

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 9, 2005, 12:50:22 ; Search time 214 Seconds  
(without alignments)  
2561.461 Million cell updates/sec

Title: US-10-724-972A-6352  
Perfect score: 1690  
Sequence: 1 GVESVRLGKILSVIGLLFVL.....EEMAKVELSKDKKDK 335

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 81818359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool\_h/US10724972/runat\_03112005\_141901\_29159/app\_query.fasta\_1.519  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40  
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10724972@cgn2\_1\_69@runat\_03112005\_141901\_29159 -NCFU=6 -ICPU=3  
-NO\_MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	475	28.1	668	4	US-08-956-171E-355
2	475	28.1	668	4	US-08-781-986A-355
3	405	24.0	999	4	US-09-830-217-15
4	405	24.0	999	4	US-10-278-946-15
5	405	24.0	3775	4	US-08-956-171E-238
6	405	24.0	3775	4	US-08-781-986A-238
7	334	19.8	2115	4	US-08-956-171E-604
8	334	19.8	2115	4	US-08-781-986A-604
9	333	19.7	242	4	US-08-956-171E-2556
10	333	19.7	242	4	US-08-781-986A-2556
11	307.5	18.2	1059	4	US-09-107-532A-862
12	296.5	17.5	1056	3	US-09-134-001C-2655

C	13	296.5	17.5	3153	4	US-09-710-279-3348	Sequence 3348, Ap
	14	296.5	17.5	3267	4	US-09-710-279-4042	Sequence 4042, Ap
	15	296.5	17.5	3618	4	US-09-710-279-3564	Sequence 3564, Ap
	16	296	17.5	1056	4	US-09-543-681A-2690	Sequence 2690, Ap
	17	292	17.3	1044	4	US-09-710-279-461	Sequence 461, Ap
	18	292	17.3	1044	4	US-09-710-279-1267	Sequence 1267, Ap
	19	291	17.2	702	4	US-09-902-540-2208	Sequence 2208, Ap
	20	291	17.2	709	4	US-09-902-540-1583	Sequence 1583, Ap
	21	283	16.7	975	3	US-09-071-035-37	Sequence 37, Appl
	22	283	16.7	984	4	US-09-134-000C-937	Sequence 937, Ap
	23	278.5	16.5	978	4	US-09-543-681A-1393	Sequence 1393, Ap
	24	276.5	16.4	954	4	US-09-891-641-16	Sequence 16, Appl
	25	268	15.9	889	3	US-09-071-035-39	Sequence 39, Appl
	26	266	15.7	1041	3	US-09-602-787A-571	Sequence 571, Appl
	27	265	15.7	183	3	US-09-134-001C-2023	Sequence 2023, Ap
C	28	257	15.2	28194	4	US-09-902-540-1250	Sequence 1250, Ap
	29	254	15.0	1035	4	US-09-583-110-1062	Sequence 1062, Ap
C	30	253	15.0	9707	3	US-08-961-521-164	Sequence 164, Ap
	31	250	14.8	966	4	US-09-107-433-2206	Sequence 2206, Ap
	32	247.5	14.6	1119	4	US-09-602-787A-597	Sequence 597, Ap
	33	241	14.3	975	4	US-09-328-352-318	Sequence 318, Appl
	34	240	14.2	954	3	US-09-071-035-69	Sequence 69, Appl
	35	238	14.1	837	4	US-09-902-540-5089	Sequence 5089, Ap
	36	236	14.0	895	3	US-08-961-083-23	Sequence 23, Appl
	37	236	14.0	895	4	US-09-536-784-23	Sequence 2609, Ap
	38	236	14.0	906	4	US-09-134-000C-2609	Sequence 71, Appl
	39	228	13.5	868	3	US-09-071-035-71	Sequence 11, Appl
	40	218	12.9	966	4	US-09-302-626B-11	Sequence 9, Appl
	41	217	12.8	966	4	US-09-107-433-1916	Sequence 1916, Ap
	42	217	12.8	966	4	US-09-107-433-1916	Sequence 1887, Ap
	43	210	12.4	900	4	US-09-710-279-1887	Sequence 2830, Ap
	44	205.5	12.2	930	3	US-09-134-001C-2830	
	45	205.5	12.2	930	3	US-09-134-001C-2830	

ALIGNMENTS

RESULT 1

US-08-956-171E-355

; Sequence 355, Application US/08956171E

; Patent No. 6593114

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

Gil H. Choi

Patrick S. Dillon

Craig A. Rosen

Steven C. Barash

Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5256

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956,171E

FILING DATE: 20-Oct-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,861

FILING DATE: January 5, 1996

APPLICATION NUMBER: 08/781,986

FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Mark J. Hyman

BEST AVAILABLE COPY

APPLICATION NUMBER: US/08/781,986A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Benson, Bob  
REGISTRATION NUMBER: 30,446  
REFERENCE/DOCKET NUMBER: PB248PP  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 355:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 668 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-781-986A-355

Alignment Scores:  
Pred. No.: 3,76e-45 Length: 668  
Score: 475.00 Matches: 91  
Percent Similarity: 86.89% Conservatives: 15  
Best Local Similarity: 74.59% Mismatches: 16  
Query Match: 28.11% Indels: 0  
Gaps: 4

US-10-724-972A-6352 (1-335) x US-08-781-986A-355 (1-668)

Qy 210 LeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLys 229  
Db 2 TTATTAGCACATCCAACTATTTCATATGTTGGACAAATTTTAAACGAACCTAGGATTTAAA 61

Qy 230 GluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeu 249  
Db 62 AATGCAATTAAGTACGATGTAACAAAGGTTTAAAGTAAATTTTAAAGGACCTTACTTA 121

Qy 250 GlnMetAenThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn 269  
Db 122 CAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAGCGTATGATCATTTATGACAGAT 181

Qy 270 LysAlaSerSerAsnGluProSerLeuLysGluLeuGluLysAspProValTirpLysLys 289  
Db 182 CATGCTAAAAAGATTCCTGCTGAATTCAGAAAGTTTACAAAGAGATGCAACATGGAAAAAG 241

Qy 290 LeuAenAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTirpAlaArgSer 309  
Db 242 TTGAATGCAGTTAAAAAATAATCGCGTGGATATTGTTGACCGTGATGTTGGCAAGATCT 301

Qy 310 ArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAsp 329  
Db 302 CGTGCCTTAATTTCTTCTGAAGAAATGCTTAAAGAACTTGTGTAATTATCAAAAAAGAA 361

Qy 330 SerLys 331  
Db 362 CAAAAG 367

RESULT 3  
US-09-830-217-15  
Sequence 15, Application US/09830217  
Patent No. 6521441  
GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides  
FILE REFERENCE: PB461PCT  
CURRENT APPLICATION NUMBER: US/09/830,217  
CURRENT FILING DATE: 2001-04-24  
PRIOR APPLICATION NUMBER: PCT/US99/06199  
PRIOR FILING DATE: 1999-03-18  
PRIOR APPLICATION NUMBER: 60/080,296  
PRIOR FILING DATE: 1998-04-01

REGISTRATION NUMBER: 46,789  
REFERENCE/DOCKET NUMBER: PB248P1  
TELEPHONE: (240) 314-1224  
TELEFAX: (301) 309-8439  
INFORMATION FOR SEQ ID NO: 355:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 668 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-956-171E-355

Alignment Scores:  
Pred. No.: 3,76e-45 Length: 668  
Score: 475.00 Matches: 91  
Percent Similarity: 86.89% Conservatives: 15  
Best Local Similarity: 74.59% Mismatches: 16  
Query Match: 28.11% Indels: 0  
Gaps: 4

US-10-724-972A-6352 (1-335) x US-08-956-171E-355 (1-668)

Qy 210 LeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLys 229  
Db 2 TTATTAGCACATCCAACTATTTCATATGTTGGACAAATTTTAAACGAACCTAGGATTTAAA 61

Qy 230 GluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeu 249  
Db 62 AATGCAATTAAGTACGATGTAACAAAGGTTTAAAGTAAATTTTAAAGGACCTTACTTA 121

Qy 250 GlnMetAenThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn 269  
Db 122 CAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAGCGTATGATCATTTATGACAGAT 181

Qy 270 LysAlaSerSerAsnGluProSerLeuLysGluLeuGluLysAspProValTirpLysLys 289  
Db 182 CATGCTAAAAAGATTCCTGCTGAATTCAGAAAGTTTACAAAGAGATGCAACATGGAAAAAG 241

Qy 290 LeuAenAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTirpAlaArgSer 309  
Db 242 TTGAATGCAGTTAAAAAATAATCGCGTGGATATTGTTGACCGTGATGTTGGCAAGATCT 301

Qy 310 ArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAsp 329  
Db 302 CGTGCCTTAATTTCTTCTGAAGAAATGCTTAAAGAACTTGTGTAATTATCAAAAAAGAA 361

Qy 330 SerLys 331  
Db 362 CAAAAG 367

RESULT 2  
US-08-781-986A-355  
Sequence 355, Application US/08781986A  
Patent No. 6737248  
GENERAL INFORMATION:  
APPLICANT: Charles Kunsch  
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 5255  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:

```

; PRIOR APPLICATION NUMBER: 60/084,674
;
; PRIOR FILING DATE: 1998-05-07
;
; NUMBER OF SEQ ID NOS: 22
;
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 15
;
; LENGTH: 999
;
; TYPE: DNA
;
; ORGANISM: Staphylococcus aureus
US-09-830-217-15

```

Alignment Scores:		
Pred. No.:	7,578-37	999
Score:	405.00	Matches: 113
Percent Smilarity:	51.74%	Conservative: 65
Best Local Smilarity:	32.85%	Mismatches: 132
Query Match:	23.96%	Indels: 34
DB:	4	Gaps: 12

US-10-724-972A-6352 (1-335) x US-09-830-217-15 (1-999)

Qy	18	Leu	Leu	Ile	Leu	Ser	Val	Ile	Gly	Leu	Leu	Phe	Val	Leu	Ile	Ala	Thr	Ala	Ala	Cys	Gly	27	
																					27		
Db	19	ATT	AAA	AAT	GCTT	---	GTT	GTT	ACG	CTT	GCTT	CTT	CTT	CTT	CTT	CTT	CTT	CTT	CTT	CTT	72		
Qy	28	Asn	Asn	Ser	Ser	Ser	Asn	Ser	Ser	Gly	Ser	Gly	Ser	Ser	Lys	Asp	Gly	Val	Glu	Ile	His	47	
																					47		
Db	73	GGG	AAATTC	CAAT	AAAT	CAAT	CAT	CTG	TAT	CA	CA	CA	AA	GAT	AAG	AT	AAG	CA	CA	CA	AT	132	
Qy	48	Glu	Glu	Gly	Thr	Thr	Thr	Val	Pro	Lys	His	Pro	Lys	Arg	Val	Val	Val	Leu	Glu	Tyr	Ser	67	
																					67		
Db	133	GCA	ATGGG	TAC	AACT	TGA	AAAT	TTAA	AGG	GA	CA	CA	CA	AA	CGG	TGT	TGT	TAC	G	TAT	CA	192	
Qy	68	Phe	Val	Asp	Ala	Leu	Val	Ala	Leu	Asp	Val	Lys	Pro	Val	Gly	Ile	Ala	Asp	Asp	Leu	Lys	87	
																					87		
Db	193	GCC	ACT	G	AGC	TG	CTG	TAT	CTT	TAG	GTG	TAA	CC	TGT	AGG	TG	CTG	TGA	AT	CA	TG	252	
Qy	88	Lys	Asn	Arg	G	Ile	Ile	Lys	Pro	Leu	Asp	Lys	Ile	Gly	Lys	Tyr	Thr	Ser	Val	Gly	Thr	107	
																					107		
Db	253	CAAAA	ACC	CGAA	ATT	CGA	AT	TAC	ATA	AAAA	TGAT	TTAA	AA	GAT	ACT	AAG	AT	TG	TG	AT	CGA	312	
Qy	108	Arg	Lys	Gln	Pro	Asn	Leu	Glu	Glu	Ile	Ser	Lys	Leu	Lys	Pro	Asp	Leu	Ile	Ala	Asp	127		
																					127		
Db	313	GAAC	CTG	CAC	CTA	ACT	TAG	AGG	AAAT	CTCT	AAAT	TTAA	AA	CCG	GAC	TCT	TAAT	TG	TCG	CGT	CA	372	
Qy	128	Asn	Asn	Arg	His	Lys	Gly	Ile	Tyr	Lys	Asp	Leu	Asn	Lys	Ile	Ala	Pro	Thr	Ile	Glu	Leu	147	
																					147		
Db	373	AAAG	TTAG	AAAT	TGA	AAAA	AGTTT	TAC	GAT	CA	AT	TAT	CTA	AAAA	TG	CGA	CA	CA	CA	CA	AGTT	426	
Qy	148	Lys	Ser	Phe	Asp	Gly	Asp	Tyr	Asn	Glu	Asn	Ile	Asp	Ala	Phe	Lys	Thr	Ile	Ser	Lys	Ala	167	
																					167		
Db	427	---	TCT	ACT	GAT	CAT	AGTTT	TT	CAAA	TT	CAAA	---	GAT	CA	CA	CA	CTA	AGT	TTA	AT	GGG	AAAGCT	480
Qy	168	Leu	Gly	Lys	Glu	Glu	Glu	Gly	Lys	Arg	Leu	Glu	His	Asp	Lys	Lys	Ile	Glu	Leu	187			
																					187		
Db	481	TTAG	GGA	AA	AGAA	AA	AGAC	TGA	AGAT	TT	ACT												

## RESULT 4

```

US-10-278-946-15
; Sequence 15, Application US/10278946
; Patent No. 6821754
; GENERAL INFORMATION:
; APPLICANT: Simpson et al.
; TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB461USD1
; CURRENT APPLICATION NUMBER: US/10/278,946
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: US 09/830,217
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: PCT/US99/061199
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: US 60/078,682
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 60/080,296
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: US 60/084,674
; PRIOR FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 999
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-278-946-15

```

Alignment Scores:

Pred. No.:	7.576-37	Length:	999
Score:	405.00	Matches:	113
Percent Similarity:	51.74%	Conservative:	65
Best Local Similarity:	32.8%	Mismatches:	132
Query Match:	23.9%	Indels:	34
DB:	4	Gaps:	12

US-10-724-972A-6352 (1-335) x US-10-278-946-15 (1-999)

Qy	8	LeuLysIleIeuSerValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCysGly	27
Db	19	ATTAATAATGCTT--GTTGTTACGCTTGCTTCTCTACTTGT--TTAGCAGGATGCTACT	72
Qy	28	AsnAsnSerSerAsnSerSerLysGluSerSerLysAspGlyValGluIleLysHis	47
Db	73	GGGATTTCAATTAACAACATCATCTGATAACAAGATAGGAAACAACCTTCAATTAAACAT	132
Qy	48	GluGluGlyThrThrLysValProLysHisProLysArgValValValLysGluIleYrSer	67
Db	133	GCAATGGGTACAACCTGAAATTAAGGGAAACCAAGCGTGTGTTACGCTATATCAAGGT	192
Qy	68	PheValAspAlaIleValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLys	87
Db	193	GCCACTGACGTGCGTGTATCTTTTAGGTGTAAACCTGTAGTGTGTGATCATGGACA	252





Db 469 ---TCTACTGATACAGTCTTTTCAAAATTCAAA---GATACAACTAAGTTAATGGGGAAGCT 522  
Qy 168 LeuGlyLysGluGluGluGlyLysArgLeuGluGluHisAspLysLysIleGluGlu 187  
Db 523 TTAGGGAAGAAAAGAAAGCTGAAGATTTACTTAAAGAGTACGATGATAAGTAGCTGCA 582  
Qy 188 TyLysLysGluLeuThrMetAspLysAsnGlnLysValLeuProAlaValAlaLys 207  
Db 583 TTCCAAAA---GATGCAAAAGCAAAAGTATAAAGATGCGCCATTGAAA 630  
Qy 208 SerGlyLeuLeuAlaHisProSerAsn-----SerTyValGly 220  
Db 631 GCTTCAGTTGTTAACTTCGCTGCTGATCATAAGAAATTTATGCTGCGATATGCTGGT 690  
Qy 221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu 240  
Db 691 GAAATCTTAATGATTTAGGATTCAAACGT-----AATAAGACTTA 732  
Qy 241 SerLysTyLysGly-----ProTyLeuGlnMetAsnThr-----GluThrLeuSer 257  
Db 733 CAATAACAAGTTGATAATCGTAAAGATATTATCCAACTTACATCTAAAGAAAGCATTCGA 792  
Qy 258 GlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSer 277  
Db 793 TTAATGAACGCTGATCATATTTTGTAGTAAATCGATCCAAATGCGAAAGATGCTGCA 852  
Qy 278 Leu-----LysGluLeuGluLysAspProValTrpLysLysLeuAsnAla 292  
Db 853 TTAGTAAAAAGACTGAAAGCGAATGAGCTTCAAGTAAAGAGTGGAAAAATTAGACGCA 912  
Qy 293 ValLysAsnGlnArgVal---AspIleLeuAspArgAspLeuTrpAlaArgSerArgGly 311  
Db 913 GTTAAAAACAACCAAGTATCTGATGATTTAGATGAAATCAGCTTGGAACTTAGCTGGCGA 972  
Qy 312 LeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAspSerLys 331  
Db 973 TATAATCTTCATTAATAAATTTATGACGATTTATATGAA-----AAGTTAAATATTGAA 1026  
Qy 332 LysAspAsnLys 335  
Db 1027 AAACAATCAAAA 1038

RESULT 6  
US-08-781-986A-238  
; Sequence 238, Application US/08781986A  
; Patent No. 6737248  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5255  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/781.986A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Benson, Bob  
; REGISTRATION NUMBER: 30,446  
; REFERENCE/DOCKET NUMBER: PB248PP

; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 238:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3775 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-781-986A-238

Alignment Scores:  
Pred. No.: 5,58e-36 Length: 3775  
Score: 405.00 Matches: 113  
Percent Similarity: 51.74% Conservative: 65  
Best Local Similarity: 32.85% Mismatches: 132  
Query Match: 23.96% Indels: 34  
DB: 4 Gaps: 12

US-10-724-972A-6352 (1-335) x US-08-781-986A-238 (1-3775)

Qy 8 LeuLysIleLeuSerValIleGlyLeuLeuPheValLeuAlaThrAlaAlaCysGly 27  
Db 61 ATTAATAATGCTT---GTTGTTACGCTTGTCTTCTTACTTGT---TTAGCAGGATGCTAGT 114  
Qy 28 AsnAsnSerSerSerAsnSerSerLysGluSerSerLysAspGlyValGluIleLysHis 47  
Db 115 GGGAAATTCAAATAAACAATCATCTGATAACAAAGATAAGGAACAACATCTCAATTTAAACAT 174  
Qy 48 GluGluGlyThrThrLysValProLysHisProLysArgValValValValLeuGluTySer 67  
Db 175 GCAATGGGTACACACTGAATTAAGGGAAACCAAGCGGTGTTGTACGCTATATCAAGT 234  
Qy 68 PheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLys 87  
Db 235 GCCACTGACGCTGCTGTTATCTTTAGTGTATAAACCCTGTAGGTGCTAGTAATCATGGACA 294  
Qy 88 LysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyThrSerValGlyThr 107  
Db 295 CAAAAACGAAATTCGAATACATAAATAAATGATTTAAAGAGTACTAAGATTGTAGTCAA 354  
Qy 108 ArgLysGlnProAsnLeuGluLysSerLysLysLeuLysProAspLeuIleIleAlaAsp 127  
Db 355 GAACCTGCACCTTAACCTTAGAGGAAATCTCTAAATTAACACCGGACTTAAATTTGTCGGTCA 414  
Qy 128 AsnAsnArgHisLysGlyIleTyLysAspLeuAsnLysIleAlaProThrIleGluLeu 147  
Db 415 AAAGTTAGAAATGAAAAAGTTTACGATCAATTAATTAATAATCGCACCAACAGTT----- 468  
Qy 148 LysSerPheAspGlyAspTyAsnGluAsnIleAspAlaPheLysThrIleSerLysAla 167  
Db 469 ---TCTACTGATACAGTTTTCAAATTCAAA---GATACAACTAAGTTAATGGGAAAGCT 522  
Qy 168 LeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAspLysIleGluGlu 187  
Db 523 TTAGGGAAGAAAAGAAAGCTGAAGATTTACTTAAAGAGTACGATGATAAGTAGCTGCA 582  
Qy 188 TyLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLys 207  
Db 583 TTCCAAAA-----GATGCAAAAGCAAAAGTATAAAGATGATGATGCCATTGAAA 630  
Qy 208 SerGlyLeuLeuAlaHisProSerAsn-----SerTyValGly 220  
Db 631 GCTTCAGTTGTTAACTTCGCTGCTGATCATAAGAAATTTATGCTGCGATATGCTGGT 690  
Qy 221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu 240  
Db 691 GAAATCTTAATGATTTAGGATTCAAACGT-----AATAAGACTTA 732  
Qy 241 SerLysTyLysGly-----ProTyLeuGlnMetAsnThr-----GluThrLeuSer 257  
Db 733 CAAAAACAGTTGATAATGTAAGATATTATCCAACTTACATCTAAAGAAAGCATTCCTCA 792

QY 258 GlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSer 277  
Db 793 TTAATGAACGCTGATCATATTTTGTAGTAAATCAGATCCAAATCGGAAGATGCTGCA 852  
QY 278 Leu-----LysGluLeuGluLysAspProValTrpLysLysLeuAenAla 292  
Db 853 TTAGTTAAAAAGACTGAAGCGAATCGACTTCAAGTAAAGAGTGGAAAAATTTAGACGCA 912  
QY 293 ValLysAsnGlnArgVal---AspIleLeuAspArgAspLeuTrpAlaArgSerArgGly 311  
Db 913 GTTAAAAACAACCAAGTATCTGATGATGATTTAGATGAAATCAGTTGGAACTTAGCTGCGGA 972  
QY 312 LeuLysSerSerGluGluMetAlaLysGluLeuValGluLysSerLysLysAspSerLys 331  
Db 973 TATAAATCTTCATTAACACTTATTGACGATTTATATGAA-----AAGTTAAATATGAA 1026  
QY 332 LysAspAsnLys 335  
Db 1027 AACAATCAAAA 1038  
RESULT 7  
US-08-956-171E-604  
; Sequence 604, Application US/08956171E  
; Patent No. 6593114  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; Gil H. Choi  
; Patrick S. Dillon  
; Craig A. Rosen  
; Steven C. Barash  
; Michael R. Fannon  
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5256  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/956,171E  
; FILING DATE: 20-Oct-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/009,861  
; FILING DATE: January 5, 1996  
; APPLICATION NUMBER: 08/781,986  
; FILING DATE: January 3, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mark J. Hymen  
; REGISTRATION NUMBER: 46,789  
; REFERENCE/DOCKET NUMBER: PB248P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (240) 314-1224  
; TELEFAX: (301) 309-8439  
; INFORMATION FOR SEQ ID NO: 604:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2115 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 604:  
US-08-956-171E-604  
Alignment Scores: 3.35e-28 Length: 2115  
Pred. No.: 334.00 Matches: 67  
Score:

Percent Similarity: 63.84% Conservative: 46  
Best Local Similarity: 37.83% Mismatches: 58  
Query Match: 19.76% Indels: 6  
DB: 4 Gaps: 3  
US-10-724-972A-6352 (1-335) x US-08-956-171E-604 (1-2115)  
QY 154 TyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGlu 173  
Db 3 TATAATGCAATATTTGAAGCATTTAAACAGTCGCTAAAGCAGTAGGCAAGAGAAAGAA 62  
QY 174 GlyLysLysArgLeuGluGluHisAspLysLysIleGluGluLysLysLysGluLysThr 193  
Db 63 GCGCAAGACGCTCTGGAAGAGCATGATAAAATATTAGCCGAGATTAGAAAGAAATTTGAA 122  
QY 194 MetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHis 213  
Db 123 CAGAGTACGTTAAATCTGCAATTCGGTATCTCAAGACGAGGTATGTTTATTAAT 182  
QY 214 ProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSer 233  
Db 183 AATGAAGATACATTTATGGGACAAATCTTAAATTTAAATGGGTATTTCAACCTGAAGTCAMA 242  
QY 234 AspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThr 253  
Db 243 AAARAAAAAACTACGCATGTTGGTGAACGCAAGGGTGGTATTTAGCCACTGACGAAAAACGGAC 362  
QY 254 GluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn---LysAlaSer 272  
Db 303 GAAGAATTCGCAATATCAATCCAAAGTTATGATTTTAGCCACTGACGAAAAACGGAC 362  
QY 273 SerAsnGluProSerLeuLysGluLeuGluLysAspPro---ValTrpLysLysLeuAsn 291  
Db 363 AAAAAATAGAACGAAATTC-----ATTGATCTCGAGTTTGGAAATCATTAATA 410  
QY 292 AlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArgGly 311  
Db 411 GCTGTGAAGATACAAAGATTTATGACGTGACCGAAATTAAGTGTGTTGAAATCAAGGGG 470  
QY 312 LeuLysSerSerGluGluMetAlaLysGluLeuValGluLysSerLysLys 328  
Db 471 ATTATCGCAAGTGAAGTATGGCAGAGAGATTAGAAAAAATTTGCAGAAAAA 521  
RESULT 8  
US-08-781-986A-604  
; Sequence 604, Application US/08781986A  
; Patent No. 6737248  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5255  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/781,986A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Benson, Bob  
; REGISTRATION NUMBER: 30,446

REFERENCE/DOCKET NUMBER: PB248PP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 604:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2115 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-781-986A-604

Alignment Scores:  
Pred. No.: 3 35e-28 Length: 2115  
Score: 334.00 Matches: 67  
Percent Similarity: 63.84% Conservative: 46  
Best Local Similarity: 37.85% Mismatches: 58  
Query Match: 19.76% Indels: 6  
DB: 4 Gaps: 3

US-10-724-972A-6352 (1-335) x US-08-781-986A-604 (1-2115)

QY 154 TyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 173  
Db 3 TATAATGCAATATTGAAGCATTTAAACAGTCGCTAAGCAGTAGGCAAGAAGAA 62  
QY 174 GlyLysLysArgLeuGluGluHisAspLysLysIleGluGlyLysGluIleThr 193  
Db 63 GCGCAGAAGCGCTCGAAAGACGATATAAATATTAGCGAGATTAGAAAGAAATGAA 122  
QY 194 MetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHis 213  
Db 123 CAGAGTAGCTTAATCTCGCATTTGCGTATCTCAAGACGAGTATGTTTATTAAT 182  
QY 214 ProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSer 233  
Db 183 AATGAAGATACATTTATGGACATTTCTTAATTAATGGGTATTCACCTGAAGTCAMA 242  
QY 234 AspAspValThrLysGlyLysSerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThr 253  
Db 243 AAARMAAACTACGCGATGTTGGTGAACGCAAGGGTGGTCTTATATATATTAAATAT 302  
QY 254 GluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn---LysAlaSer 272  
Db 303 GAAGAACTTGGCAATATCAATCCAAAGTTATGATTTAGCCACTGACGGAACCGGAC 362  
QY 273 SerAsnGluProSerLeuLysGluLeuGlyLysAspPro---ValTyrLysLysLeuAsn 291  
Db 363 AAAAATAGAACGAATTC-----ATTGATCTCGCAGTTTGGAAATCATTAATA 410  
QY 292 AlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTyrAlaArgSerArgGly 311  
Db 411 GCTGTGAAGATAACAAAGTTTATGACGTTGACCGAATAAGTGGTGGAAATCAAGGGGG 470  
QY 312 LeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328  
Db 471 ATTATCGCAAGTGAAGTATGCGAGAAGATTTAGAAAAAATTCAGAAAAA 521

RESULT 9

US-08-956-171E-2556  
Sequence 2556, Application US/08956171E  
Patent No. 6593114

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

Gil H. Choi

Patrick S. Dillon

Craig A. Rosen

Steven C. Barash

Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5256

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/956,171E  
FILING DATE: 20-Oct-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/009,861  
FILING DATE: January 5, 1996  
APPLICATION NUMBER: 08/781,986  
FILING DATE: January 3, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mark J. Hyman  
REGISTRATION NUMBER: 46,789  
REFERENCE/DOCKET NUMBER: PB248P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (240) 314-1224  
TELEFAX: (301) 309-8439  
INFORMATION FOR SEQ ID NO: 2556:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 242 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 2556:  
US-08-956-171E-2556

Alignment Scores:  
Pred. No.: 1 68e-29 Length: 242  
Score: 333.00 Matches: 66  
Percent Similarity: 90.00% Conservative: 6  
Best Local Similarity: 82.50% Mismatches: 8  
Query Match: 19.70% Indels: 0  
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-08-956-171E-2556 (1-242)

QY 70 AspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsnLysLysAsn 89  
Db 2 GATGATTTAGCAGCATTAGACGTTAAACAGTTGGTATTGCTGATGATGCTAAGAAAAA 61  
QY 90 ArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLys 109  
Db 62 CGTATCATTAACACGATTAGAGAAAAAATTCGGGATTATCTCTCTGTAGGTACAGCTAAA 121  
QY 110 GlnProAsnLeuGluGluIleSerLysLysLysProAspLeuIleIleAlaAspAsnAsn 129  
Db 122 CAGCCAACTTAGAAGAAATTTAGTAATTAACCGGATTTAATATTCGCTGATAGCAGT 181  
QY 130 ArgHisLysGlyIleTyrLysAspLysLysIleAlaProThrIleGluLeuLysSer 149  
Db 182 AGACATANAGGTATTATAAAGAAATTAACCAAAATTCACCAACATTTATCATTAAAGAGT 241

RESULT 10

US-08-781-986A-2556

Sequence 2556, Application US/08781986A

Patent No. 6737248

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5255

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/781.986A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Benson, Bob  
REGISTRATION NUMBER: 30,446  
REFERENCE/DOCKET NUMBER: PB248PP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 2556:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 242 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-781-986A-2556  
Alignment Scores:  
Pred. No.: 1.68e-29 Length: 242  
Score: 333.00 Matches: 66  
Percent Similarity: 90.00% Conservative: 6  
Best Local Similarity: 82.50% Mismatches: 8  
Query Match: 19.70% Indels: 0  
Gaps: 0  
US-10-724-972A-6352 (1-335) x US-08-781-986A-2556 (1-242)  
QY 70 AspAlaLeuValAlaLeuAspValLysProValGlyLeuAlaAspAsnLysLysAsn 89  
Db 2 GATGATTAGAGCATTAGACGTTAAACCGTTGGTATTGCTGATGGTAAAGAAAAA 61  
QY 90 ArgIleLeuProLeuArgAspLysLeuGlyLysThrSerValGlyThrArgLys 109  
Db 62 CGTATCAATTAAACCGATTAGACGTTAAACCGTTGGTATTGCTGATGGTAAAGAAAAA 121  
QY 110 GlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsnAsn 129  
Db 122 CAGCCAACTTGAAGAAATTAGTAAATTAAACCGGATTTATATCGCTGATGACAGT 181  
QY 130 ArgHisLysGlyIleThrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSer 149  
Db 182 AGACATANAGGTATTATTAAGAAATTAAACAAATTAACAAATTAACAAATTAACAAATTA 241  
RESULT 11  
US-09-107-532A-862  
Sequence 862, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneka  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 862:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1059 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...1059  
SEQUENCE DESCRIPTION: SEQ ID NO: 862:  
US-09-107-532A-862  
Alignment Scores:  
Pred. No.: 1.31e-25 Length: 1059  
Score: 307.50 Matches: 107  
Percent Similarity: 50.29% Conservative: 65  
Best Local Similarity: 31.29% Mismatches: 123  
Query Match: 18.20% Indels: 47  
Gaps: 15  
US-10-724-972A-6352 (1-335) x US-09-107-532A-862 (1-1059)  
QY 5 ValargGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeuIleAlaThrAla 24  
Db 118 GTAGTAGGTATCTTGATTTCTAGCATCATTTGGTTTA-----GCA 156  
QY 25 AlaCysGly-----AsnAsnSerSerSerSerSerSerSerSerSerSerSerSer 39  
Db 157 GCATGTGGAATAGCAACACACCAGTCAGGCAGACACAAACAAAGCAGCACACAAACT 216  
QY 40 LysAspGlyValGluIleLysHisGluGluGly---ThrThrLysValProLysHisPro 58  
Db 217 GAAACTACATTCAGCATTAACAGATAGTAAATCGAGACCAAAATCGAGGTGCCAAATAATCCA 276  
QY 59 LysArgValValValLeuGluTyThrPheValAspAlaLeuValAlaLeuAspVal--- 77  
Db 277 GAAAGAGTCGTGATTCGCAATGTTCTTAGATAGATGATGCTCTTGGTGTGGA 336  
QY 78 ---LysProValGlyIleAlaAspAsnLysLysAsnArgIleIleLysProLeuArg 96  
Db 337 GATAGGTAGTTGGAGCAGCTACAAGCAGC-----CTTCTGAGTATCTT 381  
QY 97 AspLysIleGlyLysThrSerValGlyThrArgLysGlnProAsnLeuGluGluIle 116  
Db 382 TCATCTCTATAAAAAAGTAGAATCAGCAGGTGGAATAAAGAACACAGACTTAGAAAAAT 441  
QY 117 SerLysLeuLysProAspLeuIleAlaAspAsnAsnAsnHisLysGlyIleTyLys 136  
Db 442 AATCAGCTTCAGCCAGATCTAATCATCATTT---TCTGCCGACAAAGAGATTTTCAAAA 498

DB: 3 Gaps: 16

US-10-724-972A-6352 (1-335) x US-09-134-001C-2655 (1-1056)

```

Qy 1 GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeu 20
    ||||| :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 1 GGAGTTCACATATGAAAAAACAGCTCTATTATTTATTATTGTCTCTAGTTTAGTTTGA 60

Qy 21 IleAlaThrAlaLaCyseGlyAsnAsnSerSerSerAsnSerSerLysGluSer--Ser 39
    |||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 61 ACGGCTTGTAAGTAGTTCGAATAATAATTCAACTTCGAAAAAGAAAATAAGTAGTCTCT 120

Qy 40 LysaspGlyValGluIleLysHisGlu----- 48
    ||||| :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 121 AAAGAAACTGTAAACCATTCAAATAATAGTTTGAAGCAAGTGGTAAAGAAAATAATGCCAGT 180

Qy 49 -----GluGlyThrThryLysValProLysHisProLysArgValValVal 63
    ||||| :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 181 GATAAGAAAAAATCTCTAATCTGTGGAAGTACCAAGAATCTTAAAAATGCCGTTGTA 240

Qy 64 LeuGluTySerPheValAspAlaLeuValAlaLeuaspValLysProValGlyIleAla 83
    ||||| :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 241 TTAGATTATGGAGCGCTTGATGTGTG-----AAAGAATTAGGTGGGCT 285

Qy 84 AspaspLenLys-----LyslenArgIlelleLysProLeuArgasp 97
    ||||| :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 286 GATAAAGTAAAGGTTTTACCTAAAGGTCAAAATAACCAATCTTTACCTAAATTTTAGAT 345

Qy 98 -----LysIleGlyLysTyThrSerValGlyThrArgLysGlnProAsnLeuGluGlu 115
    ||||| :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 346 GAATTTAAGATGATAGTATTAATTACTCGAAATTTAAAGAAGTCAACTTTTGATANA 405

Qy 116 IleserLysLeuLysProaspLeuIle---IleAlaaspAsnAsnA-gHisLysGlyIle 134
    ||||| :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 406 GTTCATCAGCTAAACCAGATGTGATTTTATTCAGGAAGACAGCTAATCAGAAAAAT 465

Qy 135 TyrlsaspLeuasnLysIleAlaProThrIleGluleuLysSerPheaspGlyAsptyr 154
    ||||| :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 466 TTAGATGAATTTAAAAAGCTGCACCAAAAAGCTAAAGTT--GTATGTAGGTACAAGT 522

Qy 155 AsnGluAsnIle----AspAlaPheLys-----ThrIleSerLysAlaLeuGly 169
    ||||| :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 523 GATGACNACTTAATTAAGATATGAAAAAATAACAGAAAATTTAGGGAATAATCTCCAGT 582

Qy 170 LysGluGluGluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGlyTyLys 189
    ||||| :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 583 AAAGAAGATAAAGCTAAAAAATAATTAAGATTTAGATAGAAAATAATCTGATATGAAA 642

Qy 190 LysGluIleThrMethaspLysAsnGlnLysValLeuProAlaValAlaLysSerGly 209
    ||||| :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 643 GATAAAA---ACTAAAGACTTTTAATAAGAAAGTAAATGTATTATTGGTTAACGAAGGTGAA 699

Qy 210 LeuLeuAlaHis---ProserAsnSerTyValGlyGlnPheLeuSerGlnLeuGlyPhe 228
    ||||| :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 700 CTATCAAGTTTGGACGAGGAGAGATTGGGGTTAGTGGTTGTGATACATTAGGATTT 759

Qy 229 LysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyThrLeuLysGlyProTy 248
    ||||| :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 760 AAACCTGCA-----GACAAAAAGTTAGCAAAAACCCCGCAT 795

Qy 249 LeuGln---MetAsnThrGluThrLeuSerGlnValAsnPro----- 261
    ||||| :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 796 GGTCAAAATATAAAATAATGAATATATAACACAGCAGATCCAGATGTTTTTAGCTATG 855

Qy 262 GluAtqMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeu 281
    ||||| :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 856 GATCGGTGTTAGTTGTAGGTGGTAAAGCAACAACAAAT-----CAAGTT 900

Qy 282 GluLysAspProValTrpLysLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeu 301
    ||||| :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 901 TTAATAAACAAAGTTATAAAAAATGTAAAAGCAGTAAAAGTAATCATATTTACGAATTA 960

Qy 302 AspArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlalysGlu 321

```

Alignment Scores:	
Pred. No.:	2.39e-24
Score:	296.50
Percent Similarity:	47.54%
Best Local Similarity:	28.42%
Query Match:	17.54%
Length:	1056
Matches:	104
Conservative:	70
Mismatches:	131
Indels:	61

Db	961	GATCCAAACTATGGTATTCTCTTCAGGA-----TCTTCAACGACCAACTATCAACAA	1014
Qy	322	LeuValGluLeuSerLys 327	
Db	1015	ATTGATGAATTAATGAA 1032	
RESULT 13			
US-09-710-279-3348/c			
; Sequence 3348, Application US/09710279			
; Patent No. 6703492			
; GENERAL INFORMATION:			
; APPLICANT: KIMMERLY, WILLIAM JOHN			
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS			
; FILE REFERENCE: P03480US			
; CURRENT FILING DATE: 2000-11-09			
; PRIOR APPLICATION NUMBER: 60/164,258			
; PRIOR FILING DATE: 1999-11-09			
; NUMBER OF SEQ ID NOS: 4472			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 3348			
; LENGTH: 3153			
; TYPE: DNA			
; ORGANISM: Artificial Sequence			
; FEATURES:			
; OTHER INFORMATION: Description of Artificial Sequence: synthetic			
; OTHER INFORMATION: nucleic acid sequence			
US-09-710-279-3348			
Alignment Scores:			
Pred. No.:	1,248-23	Length:	3153
Score:	296.50	Matches:	104
Percent Similarity:	47.54%	Conservative:	70
Best Local Similarity:	28.42%	Mismatches:	131
Query Match:	17.54%	Indels:	61
DB:	4	Gaps:	16
US-10-724-972A-6352 (1-335) x US-09-710-279-3348 (1-3153)			
Qy	1	GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeu 20	
Db	1299	GGAGTTGCACATATGAAGAAAAACAGTCTTATTTTATTTATTTATTTCTCTAGTTTATTTA 1240	
Qy	21	IleAlaThrAlaAlaCysGlyAsnAsnSerSerSerAsnSerSerLysGluSer---Ser 39	
Db	1239	ACGGCTTGTAAGTAATAGTTTCGAATAATATTCACCTTCGAAAAAGAAAAATAGTGATCT 1180	
Qy	40	LysAspGlyValGluIleLysHisGlu----- 48	
Db	1179	AAAGAACTGTAAACCATCAAAATAGTTTGAAGCAGTGGTAAAGAAATATGCGAGT 1120	
Qy	49	-----GluGlyThrThrLysValProLysHisProLysArgValValVal 63	
Db	1119	GATAGAAAAAATCTCTAATACTGTCGAAGTACCAAGAACTCTAAAAATGCGTTGTA 1060	
Qy	64	LeuGluTyrSerPheValAspAlaLeuAlaLeuAspValLysProValGlyIleAla 83	
Db	1059	TTAGATTATGAGCGCTTGATGTGTG-----AAAGAATTAGTGTGGCT 1015	
Qy	84	AspAspAsnLys-----LysAsnArgIleIleLysProLeuArgAsp 97	
Db	1014	GATAAAGTAAAGGTTTACTTAAGTGAAATACCAATCTTTACTTAAATTTTAGAT 955	
Qy	98	-----LysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsnLeuGlu 115	
Db	954	GAATTTAAAGATGATAGTATATTAATCTGGAATTTTAAAGAAAGTGAACCTTTGATAAA 895	
Qy	116	IleSerLysLeuLysProAspLeuIle---IleAlaAspAsnAsnArgHisLysGlyIle 134	
Db	894	GTTCATCAGTAAACCCAGATGTGATTTTATTTTCAGGAAAGACAGCTAATCAGAAAAAT 835	
Qy	135	TyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGlyAspTyr 154	
Db	834	TTAGATGAATTTAAAAAGCTGCACCAAAAGCTAAAGTT---GTATATGTAGTACAAGT 778	
Qy	155	AsnGluAsnIle-----AspAlaPheLys-----ThrIleSerLysAlaLeuGly 169	
Db	777	GATGACAACTTAATTAAGATATGAAAAAATACAGAAAAATTTAGGAAAAATCTACGAT 718	
Qy	170	LysGluGluGlyLysArgLeuGluGluHisAspLysLysIleGluGluTyrLys 189	
Db	717	AAAGAAGATAAGCTAAAAAAATTAATAAGATTTAGATAGAAAAATATCTGTATATGAAA 658	
Qy	190	LysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGly 209	
Db	657	GATAAA---ACTAAGACCTTTAATAAGAAAGTAAATGATTTATTTGTTAAGAGGTGAA 601	
Qy	210	LeuLeuAlaHis---ProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPhe 228	
Db	600	CTATCAACGTTTGGACCGAGGAGAGATTGGTGGTTAGTTGTTGATACATTAGGATTT 541	
Qy	229	LysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyr 248	
Db	540	AAACCTGCA-----CACAAAAAGGTTAGCAAAAGCCGCAT 505	
Qy	249	LeuGln---MetAsnThrGluThrLeuSerGlnValAsnPro----- 261	
Db	504	GGTCAAAATATAATGAATATATTAACAAGCAGAAATCAAGATGTTATTTTAGCTATG 445	
Qy	262	GluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeu 281	
Db	444	GATCGTGTTCAGTTGTAGGTGTTAAGCAACAACAAAT-----CAAGTT 400	
Qy	282	GluLysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeu 301	
Db	399	TTAAAAAACAAAGTTATATAAAAAATGTAAGACAGTAAAAAGTAAATCATATTTACGAATA 340	
Qy	302	AspArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGlu 321	
Db	339	GATCCAAACTATGTTCTCTTCAGGA-----TCTTCAACGACCAACTATCAACAA 286	
Qy	322	LeuValGluLeuSerLys 327	
Db	285	ATTGATGAATTAATGAA 268	
RESULT 14			
US-09-710-279-4042			
; Sequence 4042, Application US/09710279			
; Patent No. 6703492			
; GENERAL INFORMATION:			
; APPLICANT: KIMMERLY, WILLIAM JOHN			
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS			
; FILE REFERENCE: P03480US			
; CURRENT APPLICATION NUMBER: US/09/710,279			
; CURRENT FILING DATE: 2000-11-09			
; PRIOR APPLICATION NUMBER: 60/164,258			
; PRIOR FILING DATE: 1999-11-09			
; NUMBER OF SEQ ID NOS: 4472			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 4042			
; LENGTH: 3267			
; TYPE: DNA			
; ORGANISM: Artificial Sequence			
; FEATURES:			
; OTHER INFORMATION: Description of Artificial Sequence: synthetic			
; OTHER INFORMATION: nucleic acid sequence			
US-09-710-279-4042			
Alignment Scores:			
Pred. No.:	1,316-23	Length:	3267
Score:	296.50	Matches:	104
Percent Similarity:	47.54%	Conservative:	70
Best Local Similarity:	28.42%	Mismatches:	131
Query Match:	17.54%	Indels:	61
DB:	4	Gaps:	16

[illegible]





```
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 461
; LENGTH: 1044
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-461

Alignment Scores:
Pred. No.: 7,74e-24 Length: 1044
Score: 292.00 Matches: 101
Percent Similarity: 47.74% Conservative: 68
Best Local Similarity: 28.53% Mismatches: 121
Query Match: 17.28% Indels: 64
DB: 4 Gaps: 16

US-10-724-972A-6352 (1-335) x US-09-710-279-461 (1-1044)
QY 16 LeuLeuPheValLeuIleAlaThrAlaCysGlyAsnAsnSerSerSerSerSer 35
DB 25 TTATTGCTCTAGTTTAAACGGCTTGTAGTAATAGTTTCAATTAATCAACT 84
QY 36 LysGlu-----SerSerLysAspGlyValGluIleLysHisGlu----- 48
DB 85 TCGAAAAGAAAATAAGTAGTCTTCAAGAACTGTACCAATCAAAATAGTTTGAAGCA 144
QY 49 -----GluclyThrThrLysValPro 55
DB 145 AGTGGTAAAGAAAATAATGCGCAGTGATAGAAAATAATCTCTAATCTCGAAGTACCA 204
QY 56 LysHisProLysArgValValLeuGluTyrSerPheValAspAlaLeuValAlaLeu 75
DB 205 AGAATCCTTAAATAATCCGCTGTGATTAGATTATGAGCGCTGATGTTG----- 255
QY 76 AspValLysProValGlyIleAlaAspAspAsnLys-----LysAsn 89
DB 256 -----AAGAATTAGTGTGCTGATAAGTAAAGTTTACCTAAAGGTGAAAATAAC 309
QY 90 ArgIleIleLysProLeuArgAsp-----LysIleGlyLysTyrThrSerValGlyThr 107
DB 310 CAATCTTTACCTAAATTTTATAGTGAATTTAAAGATGATAGTATATTAATCTGGAAT 369
QY 108 ArgLysGlnProAsnLeuGluGluLysLysLysLysLysLysLysLysLysLys 126
DB 370 TTAAAGAAAGTGAACCTTTGATAAAGTTGCATCAGCTAAACAGATGCTGATTTTATTC 429
QY 127 AspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGlu 146
DB 430 CGAAGACAGCTAATCAGAAAATTTAGATGAATTTTAAAGAGCTGCACCAAAAGCTAAA 489
QY 147 LeuLysSerPheAspGlyAspTyrAsnGluAsnIle-----AspAlaPheLys----- 162
DB 490 GTT---GTATATGTAGGTACAGTACAGTACACACTTAATTAAGATATGAAAAAATAACA 546
QY 163 ---ThrIleSerLysAlaLeuGlyLysGluGluGluGlyLysLysLysLysLysLys 181
DB 547 GAAATTTTAGGGAAATCTTACGATAAAGAAAGATAAGCTTAAATAATTAATAAGATTTA 606
QY 182 AspyLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeu 201
DB 607 GATGAAAATATCTGATGATGAAGATATAA---ACTAAAGACTTTTAAATAAGAAAGTAAATG 663
QY 202 ProAlaValAlaLysSerGlyLeuAlaHis---ProSerAsnSerTyrValGly 220
DB 664 TATTATTGGTTAACGAAGGTGAACACTCAACGTTTGGACCAAGGAGGAGGATTTGGTGGT 723
QY 221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu 240
```

```
724 TTAGTGTTCATACATTAGGATTTAAACCTGCA-----GAC 759
241 SerLysTyrLeuLysGlyProTyrLeuGln---MetAsnThrGluThrLeuSerGlnVal 259
760 AAAAAGGTTAGCAAAAGCCCGCATGTTCAAAATATAATAATGAATATAATTAACAAGCAG 819
260 AsnPro-----GluArgMetPheIleMetThrAsnLysAlaSerSer 273
820 AATCCAGATGTTATTTTAGCTATGATCGGTTCAGTTGTAGGTGCTAAGCAACAACA 879
274 AsnGluProSerLeuLysGluLeuGlyLysAspProValTyrLysLysLysLysLysAlaVal 293
880 AAT-----CAAGTTTAAAAAACAAGTTATATAAAATGTAAAGCAGTA 924
294 LysAsnGlnArgValAspIleLeuAspArgAspLeuTyrAlaArgSerArgGlyLeuIle 313
925 AAAAGTAATCATATTTACGAATTAGATCCAAACTATGTTATTTCTCTTCAGCA----- 978
314 SerSerGluGluMetAlaLysGluLeuValGluLeuSerLys 327
979 TCTTCAACGACAACTATCAACAATTTGATGAATTAATGA 1020

RESULT 18
US-09-710-279-1267
; Sequence 1267, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1267
; LENGTH: 1044
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-1267

Alignment Scores:
Pred. No.: 7,74e-24 Length: 1044
Score: 292.00 Matches: 101
Percent Similarity: 47.74% Conservative: 68
Best Local Similarity: 28.53% Mismatches: 121
Query Match: 17.28% Indels: 64
DB: 4 Gaps: 16

US-10-724-972A-6352 (1-335) x US-09-710-279-1267 (1-1044)
QY 16 LeuLeuPheValLeuIleAlaThrAlaCysGlyAsnAsnSerSerSerSerSer 35
DB 25 TTATTGCTCTAGTTTAAACGGCTTGTAGTAATAGTTTCAATTAATCAACT 84
QY 36 LysGlu-----SerSerLysAspGlyValGluIleLysHisGlu----- 48
DB 85 TCGAAAAGAAAATAAGTAGTCTTCAAGAACTGTACCAATCAAAATAGTTTGAAGCA 144
QY 49 -----GluGlyThrThrLysValPro 55
DB 145 AGTGGTAAAGAAAATAATGCGCAGTGATAGAAAATAATCTCTAATCTCGAAGTACCA 204
QY 56 LysHisProLysArgValValLeuGluTyrSerPheValAspAlaLeuValAlaLeu 75
DB 205 AGAATCCTTAAATAATCCGCTGTGATTAGATTATGAGCGCTGATGTTG----- 255
QY 76 AspValLysProValGlyIleAlaAspAspAsnLys-----LysAsn 89
DB 256 -----AAGAATTAGTGTGCTGATAAGTAAAGTTTACCTAAAGGTGAAAATAAC 309
QY 90 ArgIleIleLysProLeuArgAsp-----LysIleGlyLysTyrThrSerValGlyThr 107
DB 310 CAATCTTTACCTAAATTTTATAGTGAATTTAAAGATGATAGTATATTAATCTGGAAT 369
QY 108 ArgLysGlnProAsnLeuGluGluLysLysLysLysLysLysLysLysLysLys 126
DB 370 TTAAAGAAAGTGAACCTTTGATAAAGTTGCATCAGCTAAACAGATGCTGATTTTATTC 429
QY 127 AspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGlu 146
DB 430 CGAAGACAGCTAATCAGAAAATTTAGATGAATTTTAAAGAGCTGCACCAAAAGCTAAA 489
QY 147 LeuLysSerPheAspGlyAspTyrAsnGluAsnIle-----AspAlaPheLys----- 162
DB 490 GTT---GTATATGTAGGTACAGTACAGTACACACTTAATTAAGATATGAAAAAATAACA 546
QY 163 ---ThrIleSerLysAlaLeuGlyLysGluGluGluGlyLysLysLysLysLysLys 181
DB 547 GAAATTTTAGGGAAATCTTACGATAAAGAAAGATAAGCTTAAATAATTAATAAGATTTA 606
QY 182 AspyLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeu 201
DB 607 GATGAAAATATCTGATGATGAAGATATAA---ACTAAAGACTTTTAAATAAGAAAGTAAATG 663
QY 202 ProAlaValAlaLysSerGlyLeuAlaHis---ProSerAsnSerTyrValGly 220
DB 664 TATTATTGGTTAACGAAGGTGAACACTCAACGTTTGGACCAAGGAGGAGGATTTGGTGGT 723
QY 221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu 240
```



APPLICANT: Gil H. Choi  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 496  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071,035  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: A. Anders Brookes  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB369P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 975 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-09-071-035-37

Alignment Scores:  
Pred. No.: 5,64e-24 Length: 709  
Score: 291.00 Matches: 70  
Percent Similarity: 51.95% Conservative: 50  
Best Local Similarity: 30.30% Mismatches: 93  
Query Match: 17.22% Indels: 18  
DB: 4 Gaps: 5

US-10-724-972A-6352 (1-335) x US-09-902-540-1583 (1-709)

```
QY 104 SerValGlyThrArgLysGlnProAsnLeuGluGluLeuSerLysLysLeuLysProAspLeu 123
Db 53 AGTGTGGTACAGAAAGCGCCAGCTTAGAGCAATCGCTCTTAAAGCGCTGACTTG 112
QY 124 IleAlaAspAsnArgHisLysGlyIleTyLysAspLeuAsnLysIleAlaPro 143
Db 113 ATTATCGGAATAAATGCGACAGAAAGTATACGATCAGCTTAGTCAATTTGCTCG 172
QY 144 ThrIleGluLeuLysSerPheAspGlyAspTyAsnGluAsnIleAspAlaPheLysThr 163
Db 173 ACTGTATCTCTGAGGAGCGCGGAGATTGGAATCAAC-----TTTAGCTA 223
QY 164 IleSerLysAlaLeuGlyLysGluGluGluGlyLysArgLeuGluGluHisLeuLys 183
Db 224 TATGCAAAAGCTGTAACAAAGAGAAAGAGTAAAGAGTATTAGCTGACTATGACAAT 283
QY 184 LysIleGluGluTyLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAla 203
Db 284 CGCGTAGCAGATTAAAGAAAGCGCTTGGAGATCACTTAAACAAATCTCAGTTGTA 343
QY 204 ValAlaAlaLysSerClyLeuLeuAlaHisProSerAsnSerTyLysValGlyGlnPheLeu 223
Db 344 CGCTTTACAGCTGGAGACCTTCGTATCTATCATATAAGATTCTCAGGTGTTATTTA 403
QY 224 SerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyr 243
Db 404 GATCAGCTGGATTGCTCGCCTGAGTCTCAAGAC----- 439
QY 244 LeuLysGlyProTyLysGlnMetAsnThr-----GluThrLeuSerGlnValAsnPro 261
Db 440 --AAGGATGATTGCTGAAATGAATGCAACAAAGAACGATTCGCGCAATGACGCGC 496
QY 262 GluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeu 281
Db 497 GATCAGCTATTATTCTCTCTTATGAAACGGGAGAGCGGTGAAGCAACAAAGCTTGA 556
QY 282 Glu-----LysAspProValTrpLysLysLeuAsnAlaValLysAsnGluArgValAsp 299
Db 557 GAATGGATAACGATCTCTCTTTAAAGCTTAAAGTGGCTCAAGAAAAATCATGTTGAT 616
QY 300 IleLeuAspArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAla 319
Db 617 AAGTAGATGATGCTTACTTGGATACACACGCGGTGTTAGCAGCTAATATCGTACTT 676
QY 320 LysGluLeu-----ValGluLeuSerLysLys 328
Db 677 GATGATATTGAAAAAATCTTCTTAGATAAAAA 709
```

RESULT 21

US-09-071-035-37  
Sequence 37, Application US/09071035  
Patent No. 6448043  
GENERAL INFORMATION:

Alignment Scores:  
Pred. No.: 7,55e-23 Length: 975  
Score: 283.00 Matches: 101  
Percent Similarity: 48.19% Conservative: 59  
Best Local Similarity: 30.42% Mismatches: 130  
Query Match: 16.75% Indels: 42  
DB: 3 Gaps: 13

US-10-724-972A-6352 (1-335) x US-09-071-035-37 (1-975)

```
QY 11 LeuSerValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCysGlyAsnAsn--- 29
Db 52 GTTCAATGATGGGACTGTTAATGTTA-----AGTGTCTTGTCAACAAATAAA 99
QY 30 -----SerSerAsnSerSerLysGluSerSerLysAspGlyValGluIleLysHis 47
Db 100 AAAACAGCAGATTCTGCAACACACAGAACACACAGTAAACGGAAGTCACAGTCAAGAC 159
QY 48 GluGluGlyThrThrLysValProLysHisProLysArgValValValLeuGluTyrSer 67
Db 160 ACCAATGGTCAATTAACCGTTCCCAAAAATCTTAAGAAAGTCGTTGTTTTGATAATGTT 219
QY 68 PheValAspAlaLeuValAlaLeuAspVal-----LysProValGlyIleAlaAspAsp 85
Db 220 TCCTTGATACAATGGATGGATGCACTAGGTGCGGTGACCGCGGTGAGGTGCGCCCACTAAA 279
QY 86 AsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSerVal 105
Db 280 AAT-----ATCCCTGCGTATTGTAATAAATACCAAAAAGTTGAATCAGCA 324
QY 106 GlyThrArgLysGlnProAsnLeuGluGluLeuSerLysLysLysProAspLeuIle 125
Db 325 GCGCGCATTAAGAACCCAGATTTAGAAAAAATCAATCAACTAAAAACCAGACTTAATATT 384
QY 126 AlaAspAsnAsnArgHisLysGlyIleTyLysAspLeuAsnLysIleAlaProThrIle 145
```

```

385 ATT---TCTGGTGGTCAACAGATTATCAAGAAACAATTAAGGCAATTCGCCAACCAATT 441
146 GluLeu-----LysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPhe 161
442 TACTTACCTAGTAGATGCCAAATTCCTTGGGCATCAACAGAAACAATAATCGAAACGTTA 501
162 LysThrIleSerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHis 181
502 GGCACCTATT-----TTTGATAAAGAGAGGTAGCTAAAGAAAAAATAACTGGCTTA 552
182 AspLysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGluLysValLeu 201
553 GAAAAAGAAATTCGTGACGTGAAAAAACAAGCAGAGCTAGCGCGAAT---AATGCGCTT 609
202 ProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGln 221
610 GTTGTGTAGTTAAGGAGCAACTTTCGCTTACGGAAGAGCTCTCGTTTCGGTTTA 669
222 PheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAsp-----ValThr 237
670 ATTCATGATACATTTGGCTTCAAGCAGCA-----GACGATAAGATTGAAGCTTCCACT 723
238 LysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSer 257
724 CATGGCAAACT-----GTTTCTTACGAATATGTTTTA 756
258 GlnValAsnProGluArgMetPheIleMet---ThrAsnLysAlaSerSerAsnGluPro 276
757 GAAAAAATTCCTGGGATTCCTTTGTGGTAGATCGCACCAAGCAATTTGGTGGCAGCAT 816
277 SerLeuLysGluLeuGluLysAspProValTyrLysLysLeuAsnAlaValLysAsnGln 296
817 TCAAAAGATAGCTCGCTGCAACGAATTGATTCAAAAACCGATGCTGTGTAATAATGAT 876
297 ArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGlu 316
877 AAGTCATTATGCTTCAACAGATGCTTGTATCTAAGCGGTGGTGGATTAGAAATCAATG 936
317 GluMetAlaLysGluLeuValGluLeuSerLysLys 328
937 CATTTG-----ATGATAGAGATGTTAAAAAA 963

RESULT 22
US-09-134-000C-937
; Sequence 937, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 937
; LENGTH: 984
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-937

Alignment Scores:
Pred. No.: 7,66e-23 Length: 984
Score: 283.00 Matches: 101
Percent Similarity: 48.19% Conservative: 59
Best Local Similarity: 30.42% Mismatches: 130
Query Match: 16.75% Indels: 42
DB: 4 Gaps: 13

US-10-724-972A-6352 (1-335) x US-09-134-000C-937 (1-984)

```

RESULT 23
US-09-543-681A-1393
; Sequence 1393, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:



NAME: A. Anders Brookes  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB369P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 889 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-09-071-035-39

Alignment Scores: 3,47e-21 Length: 889  
Pred. No.: 268.00 Matches: 96  
Score: 268.00  
Percent Similarity: 47.32% Conservative: 54  
Best Local Similarity: 30.28% Mismatches: 129  
Query Match: 15.86% Indels: 38  
DB: 3 Gaps: 12

US-10-724-972A-6352 (1-335) x US-09-071-035-39 (1-889)

QY 26 CysGlyAsnAsn-----SerSerSerAsnSerSerLysGluSerSerLysAspGly 42  
Db 2 TGTCAAAACAAATAAAAAACAGCAGATCTGCACAAACAGAAACACAGCTAAACCGAA 61  
QY 43 ValGluLeuLysHisGluGluGlyThrThrLysValProLysHisProLysValVal 62  
Db 62 GTCAAGTCAAGACCAATGTCATTAACCGTCTCCCAAAATCTTAAGAAAGTCGTT 121  
QY 63 ValLeuGluThrSerPheValAlaLeuValAlaLeuAspVal-----LysProVal 80  
Db 122 GTTTTGTGATAATGTTCTCTGGATACATGATGATGATGATGATGATGATGATGAT 181  
QY 81 GlyLeuAlaAspAspAsnLysLysAsnArgLysLysLysLysLysLysLysLysLys 100  
Db 182 GGTGCGCAACTAAAAAT-----ATCCTCGCTGCTCAACAGATTTATCAAGAAACAAT 226  
QY 101 LysThrThrSerValGlyThrArgLysGlnProAsnLeuGluGluLysLysLysLys 120  
Db 227 AAGTTGAATCAGCAGCGCGCATTAAGAACACCATTTAGAAAAAATCAATCAACTAAA 286  
QY 121 ProAspLeuLeuAlaAspAsnArgHisLysGlyLysLysLysLysLysLysLysLys 140  
Db 287 CCAGACTTAATTAATTATT---TCTGTCCTCAACAGATTTATCAAGAAACAATTAAG 343  
QY 141 IleAlaProThrIleGluLeu-----LysSerPheAspGlyAspTyrAsnGlu 156  
Db 344 ATTGCGCAACCATTTACTTAGCTGTAGATGCAAAAAATCCTTGGGCATCAACGAAACA 403  
QY 157 AsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGluGlyLys 176  
Db 404 AATATCGAAACGTTAGGCATATT-----TTTGATAAAGAAAGAGGTAGCTAAAGAA 454  
QY 177 ArgLeuGluGluHisAspLysLysLysLysLysLysLysLysLysLysLysLysLys 196  
Db 455 AAAATACTGGCTTAGAAAAAGAAATTCGTGACGTGAAAAAACAACAGCAGAACTACG 514  
QY 197 AsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsn 216  
Db 515 AAT---AATGCGCTTGTGTGTAGTTAACGAAGGACAACTTTCGCTTACGGAAGAGC 571  
QY 217 SerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAsp 235  
Db 572 TCTCGTTTCGGTTTAAATTCATGATCATTTGGCTTCAAGAGCAGCA-----GACGATAAG 625  
QY 236 -----ValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsn 252  
Db 626 ATTGAAGCTTCCACTCATCGGCAAACT-----GTTTCT 658  
QY 253 ThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMet---ThrAsnLysAla 271

124 IleAlaAlaAspAsnArgHisLysGlyLysLysLysLysLysLysLysLysLysLysLys 143  
Db 355 ATCTCTGCTTCAACAAAGTTTCGGAAGAAACCGCTGCAAAAAATCAGCAGCAGCAGCAG 414  
QY 144 ThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThr 163  
Db 415 ACATCCAGCTTCTCATATCTCTTCAACTGGAAGAAACATG-----ATGCTT 465  
QY 164 IleSerLysAlaLeuGlyLysGluGluGlyLysLysLysLysLysLysLysLysLysLys 183  
Db 466 CTTGCCAGCTGAGTGGAAAGAGAGAAAGCAAGAAATTTATTCAGACTATGAACAG 525  
QY 184 LysIleGluGluTyrLysLysGluLeuThr---MetAspLysAsnGlnLysValLeuPro 202  
Db 526 GATCTAAAGAAATATAAAACAAATAACACGATAAACGGAAGATTCAAAAGCCTTGTC 585  
QY 203 AlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe 222  
Db 586 ATCAGATCAGACAGCAACATTTACATTTACCTGACACAGGTGTATTTCAACTCCACA 645  
QY 223 Leu---SerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSer 241  
Db 646 CTATACGGTGTATTAGGCTTAAAG-----GCGCGCAACGAAGTAAAGGCTGCAAAAGCG 699  
QY 242 LysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnPro 261  
Db 700 CAAGAGCTGAGT-----TCATTAGAAAAATTAAGTGAATGAATGACCCG 741  
QY 262 GluArgMetPheIle---MetThrAsnLysAlaSerSerAsnGluPro---SerLeuLys 279  
Db 742 GACCATATTTTCGTCCTCAATTTCTGATGATGAAATGCAGACAACTGATGCTTAAAA 801  
QY 280 GluLeuGluLysAspProValTrpLysLysLysLysLysLysLysLysLysLysLysLys 299  
Db 802 GATTAGAGAAAAATCCAATCTGGAAGAGCCTTAAAGCAGTCAAAAGAGACCATGTGTAT 861  
QY 300 IleLeuAspArgAsp-----LeuTrpAlaArgSerArgGlyLeu 312  
Db 862 GTCAACTCAGTGGACCTCTCGCACAGCGCGCAGCAGCTTGAGCAAGATCGCTTCTCTG 921  
QY 313 IleSerSerGluGlu 317  
Db 922 AAAGCGGCTGCTGAA 936

RESULT 25  
US-09-071-035-39  
Sequence 39, Application US/09071035  
Patent No. 6448043  
GENERAL INFORMATION:  
APPLICANT: Gil H. Choi  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 496  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071,035  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:



[illegible]

```

Db          13 CAGATTGAAAAAAGACCTGTGATGGAAAGAATAAACCCTGTGAATAATCAACCGTGTGAT 72
QY          300 ILeuAspArgAspLeuTPAlaaqSerArgGLyLeulleSerSerGluGluMetAla 319
Db          73 ATTTTAGCCGTGACTTATCGGCAAGATCACGTGTTTTTAATTTCTTCAGAAAGAAATGGCA 132
QY          320 LysGluLeuValGluLeuSerLysLysAspSerLysLysAspAsnLys 335
Db          133 AAAGAACTTTGTAATTTATCTAGAAAGATAGTAAAAAAGATAATAAG 180

RESULT 28
US-09-902-540-1250/c
; Sequence 1250, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902.540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1250
; LENGTH: 28194
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1250

Alignment Scores:
Pred. No.:      1,15e-17      Length:      28194
Score:           257.00      Matches:       82
Percent Similarity: 50.1%     Conservative:   78
Best Local Similarity: 25.71% Mismatches:    143
Query Match:        15.21% Indels:         16
DB:                  4      Gaps:            9

US-10-724-972A-6352 (1-335) x US-09-902-540-1250 (1-28194)
QY          8 LeuLysIleLeuSerValledGlyLeuLeuPheValLeuIleAlaThrAlaAlaCysGly 27
Db          23450 CTTCCTCGTGCTGCTGCTGCCGCTCTGTGGTGGGTGGGTGCAGGCCCGCGAC 23391
QY          28 AsnAnSerSerSerAsnSerSerLysHisGluSerLysAspGlyValGluIleLysHis 47
Db          23390 GCCGCGCGCTCCACTCCCCCGCGAGGCTGCCCGCGGAGGCGCGGCGGCGGACGAC 23331
QY          48 GluGluGlyThrThrllysValProLysHisProLysArgValValLeuGluTySer 67
Db          23330 GGACAAGGACACACCGTCGTCCTCCGCTGAAGCCGCGGCGGTGGTGTTCGACCTGGTG 23271
QY          68 PheValAspAlaLeuValAlaLeuAspLysProValGlyIleAlaAspAsnLys 87
Db          23270 GCGCTGGACATCTCCAGGCGCTTGGAGGTGCAGCTCCAGCGCGTGGCGGCGCATGTTC 23211
QY          88 LysAsnArgIlellellysProLeuArgAspLysIleGlyLysTyThrSerValGlyThr 107
Db          23210 CCCAGCACCTGACCCCGG---TTCCGGGAC-----CGGAAGTACCCGCGCATGGGCACC 23160
QY          108 ArgLysGlnProAsnLeuGluIleSerLysLysLysProAspLeuIleAlaAsp 127
Db          23159 CTGTTGAGCCTGACTACGAGGCACTCCAGGCCGCGAAGCGGACCTCATCATCACC--- 23103
QY          128 AsnAnArgHisLysGlyIleTyLysAspLeuAsnLysIleAlaProThrIleGluLeu 147
Db          23102 GGGGGCGCTCCAGCGCGAAGTACTCGAACCTCTCCCGCATCGCGCCACCATCGACGTG 23043
QY          148 LysSerPheAspGlyAspTyTrAsnGluAsnIleAspAla---PheLysThrIleSerLys 166

```

US-10-724-972A-6352 (1-335) x US-09-134-001C-2023 (1-183)

Qy 280 GluLeuGluLysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAsp 299

QY 148 LysSerPheAspGlyAspTyrAsnGluAsnIleAspAla---PheLysThrIleSerLys 166



Db 23042 CCGATGAGCGGCAAGGACTTCTACCGCTCGGTGATTGCCAACAGGAGATGCTAGCCAGC 22983  
Qy 167 AlaleuGlyLeuGluGluGlyLysLysArgLeuGluGluHisAspLysLeuGlu 186  
Db 22982 GTCCTTGGCAAGGAGCAAGCGCGCGCTGATTGAGGACCTGCGCAAGTCGTGCG 22923  
Qy 187 GluTyrLysLysGluLeuMetAspLysAsnGlnLysValLeuProAlaValAla 206  
Db 22922 GACCTC---CAACAGATACCGGCACACGCGC---AAGGGCTGGTCTGCTCCTCACC 22869  
Qy 207 LysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeu 226  
Db 22868 GGAGGGCGCATAGCGCTATGGCGGTTTCGGCTTGGCGTCATCCATCCGTCCTC 22809  
Qy 227 GlyPheLysGluAlaLysAspValThrLysGlyLeuSerLysTyrLysGly 246  
Db 22808 GCGTGGCGGAGGCGCTC---GAGGGCTGCGCACCTCGCTGCACGCG 22764  
Qy 247 ProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheLe 266  
Db 22763 GAG-----TCCATCAGCGGAGATTCTCCGGAGGAAGATCCCGACTGGCTGCTC 22710  
Qy 267 MetThrAsnLysAlaSerSer-----AsnGluProSerLeuLysGluLeuLysAsp 284  
Db 22709 ATCGACCGGACGCGCCACCGGTCAGAGGAAGCAATCCCGCAGGTGCTGGACAT 22650  
Qy 285 ProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspLeuLeuAspArg 304  
Db 22649 GAATTGGTGGCGAGACGCGCTGGCAGAGAACCCAGGTCTATCTACCTGGACCCGCG 22590  
Qy 305 LeuTrpAlaArgSerArgGlyLeuLeuSerSerGluGluMetAlaLysGluLeuVal 323  
Db 22589 GTCACCTACTGACCGCGCGGCTCCAGTCCGTCGAAGCAGCTCCGCGCACCGAGTGC 22533

## RESULT 29

US-09-583-110-1062  
; Sequence 1062, Application US/09583110  
; Patent No. 6699703  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al.  
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
; FILE REFERENCE: Pneumoniae for Diagnostics and Therapeutics  
; CURRENT APPLICATION NUMBER: US/09/583,110  
; CURRENT FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/107,433  
; PRIOR FILING DATE: 1998-06-30  
; PRIOR APPLICATION NUMBER: US 60/085,131  
; PRIOR FILING DATE: 1998-05-12  
; PRIOR APPLICATION NUMBER: US 60/051,553  
; PRIOR FILING DATE: 1997-07-02  
; NUMBER OF SEQ ID NOS: 5322  
; SEQ ID NO 1062  
; LENGTH: 1035  
; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
US-09-583-110-1062

Alignment Scores:  
Pred. No.: 1,77e-19 Length: 1035  
Score: 254.00 Matches: 93  
Percent Similarity: 46.56% Conservative: 56  
Best Local Similarity: 29.06% Mismatches: 119  
Query Match: 15.03% Indels: 52  
DB: 4 Gaps: 14

US-10-724-972A-6352 (1-335) x US-09-583-110-1062 (1-1035)

Qy 13 ValLeuGlyLeuPheValLeuLeuAlaThrAlaAlaCysGlyAsnAsnSer----- 30  
Db 28 CTAATAGCTATTTAGCTATGCTATAGTATGTTTACGGCTTGTCTTCAATCTGTTAA 87  
Qy 31 ---SerSerAsnSerSerLysGluSerSerLysAspGlyValGluLeuLysHisGluGlu 49

Db 88 AATGAAGAAATACTTCTTAAGAGCATGCGCTGATAAAATAGTTTAGATCATGCTTTC 147  
Qy 50 GlyThrThrLysValProLysHisProLysArgValValValLysGluTyrSerPheVal 69  
Db 148 GGTCAAACTATTATTAGATAAAAAACCTCAAGAGCTTCAACTATTGCTGGGAAATCAT 207  
Qy 70 AspAlaLeuValAlaLeuAspValLysProValGlyLeuAlaAspAsn----- 86  
Db 208 GATGTACCATAGCTTTAGGAATAGTCTCTGTTGGATTTTCAAAAGCAAAATTTACGGTGA 267  
Qy 87 LysLysAsnArgLeuLeuLysPro-----LeuArgAspLysLysLeuGlyLys 101  
Db 268 AGTGCTGATAAAGGAGTTTTACCATGACAGCAAGAAAAATCAAGAACTAAATGGTAAA 327  
Qy 102 TyrThrSerValGlyThrArgLysGlnProAsnLeuGluLeuSerLysLeuLysPro 121  
Db 328 GCTAACCTATTTCACCATTTGGATGGACTTAACTTGAAGCAATATCAAAATTTCTAAACCA 387  
Qy 122 AspLeuLeuLeuAlaAspAsnAsnArgHisLysGlyLe-----TyrLysAsp 137  
Db 388 GATGTTATCTTACAGGT-----TATTCGGTATACTAAAGAAAGATTATGACACT 438  
Qy 138 LeuAsnLysLeuAlaProThrThrLeuLeuLysSerPheAspGlyAspTyrAsnGluAsn 157  
Db 439 CTATCAAAATTCCTCTCTAGCAGCATCAAAATCTAAACCTTGGCAAACTTTATGGAGA 498  
Qy 158 IleAspAlaPheLysThrLysSerLysAlaLeuGlyLysGluGluGly----- 174  
Db 499 ---GATATGATTTAAATTCATCAAAAGCCTTAGGTATGAAAAAGAGGTGATGAGTTA 555  
Qy 175 -----LysLysArgLeuGluGluHisAspLysLysLeuGlu 186  
Db 556 ATCAAAATACTGAAGCTCGTATATCCAAAGAAATTAGAAAAACAT-----CCA 603  
Qy 187 GluTyrLysLysGluLeuThrMetAspLysAsnGlnLysValLeuProAlaVal----- 204  
Db 604 GAAATCAAGGAAGAAATC-----AAGGAAAAAAGATTTATTACTATGATTAAT 654  
Qy 205 ---AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeu 223  
Db 655 GCTGCAGATACATCAAAATTCGTGATTATATAGCAAGATCCAAAGACCAAAATTTATTA 714  
Qy 224 SerGlnLeuGly-----PheLysGluAlaLeuSerAspValThrLysGlyLeuSer 241  
Db 715 ACAGATTAGGTCTAGTTTCTCCTGAATCATTTAAAGAAATTTGAGAGTGAA-----GAT 768  
Qy 242 LysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnPro 261  
Db 769 AGTTTTCGAAG-----GAAATTTCTGAGAGAAAGCAAAATAAGATAATGAT 816  
Qy 262 GluArgMetPheLeuMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeu 281  
Db 817 GCTGATGTAATCATV-----ACTTATGCTGATGATAAAACTCTTGAAGCTTTA 864  
Qy 282 GluLysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspLeuLeu 301  
Db 865 CAAAGAGATCTCTTTTAGGTAAAAATAATGCAATTTAAATAATGGTGGCGTCTGCTAATT 924

## RESULT 30

US-08-961-527-164/c  
; Sequence 164, Application US/08961527  
; Patent No. 6420135  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA



Db 109 TCTGCTCCACAGAGGTAACCATTAAGAGTTCACTGGACGAGGTCAAACTTTCCAAAGTT 168  
Qy 58 ProLysArgValValLeuGluTyrSerPheValAlaLeuValAlaLeuAspVal 77  
Db 169 CCTGAAAGATTGTGACCTTGGACCTCGCGCTCGGATACCTATTCGCGCTTGGAGTTT 228  
Qy 78 LysProValGlyIleAlaAspAsnLysLysAsnArgIleIleLysPro----- 94  
Db 229 -----GAAAAAATATCGTGGAAATCGCTTACAAAAACT 261  
Qy 95 -----LeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGln 110  
Db 262 GTTCCGACTTATCTAAAGACCTAGTGGAACTGCTCAAAATGTGTGTTTATGAAAGAA 321  
Qy 111 ProAsnLeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsnArg 130  
Db 322 CCTGATTAGAGCTATCGCGCTTGGACCTGAGCTGATTGATTGCTTCCACACGTACA 381  
Qy 131 HisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLysSerPhe 150  
Db 382 CAAAAATTCGTAGACAAA-----TTCAAGAATAATCGCCCAACCGTTCTTCCAAAGCAAGC 438  
Qy 151 AspGlyAspTyrAsnGluAsnIleAspAla-----PheLysThrIleSerLysAlaLeuGly 169  
Db 439 AAGGACGACTACTGGACTTCTACCAAGGCTAATATCGAATCCTTAGCAAGTGCCTTCGGC 498  
Qy 170 Lys-----GluGluGlyLysArgLeuGluHisAspLysLysIleGluGlu 187  
Db 499 GAAACTGGTACACAGAAAGCAAGGAATTGGCCAGCTAGCAGACGATCCAGAA 558  
Qy 188 Tyr-----LysLysGluIleThrMetAspLysAsnGlnLys 199  
Db 559 GTGCGTACTAAATAAGAGCTCTGCAAAAAGCGCTTGGCATCTCTCTTAATGAAGAA 618  
Qy 200 ValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrVal 219  
Db 619 AAAATGGCAGCTTGTGTCGCAATCTCGTTTC----- 651  
Qy 220 GlyGlnPheLysSerGln-----LeuGlyPheLysGluAlaLeuSerAspValThrLys 238  
Db 652 ---TCTTCTTGTACCAACCTTGAAATTCAAACCACT-----GATCAACATTT 699  
Qy 239 GlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGln 258  
Db 700 GAAGACTCAGCCAGCGACAA-----GAAGTCAGCTTGAAGTGTCAAGAA 747  
Qy 259 ValAsnProGluArgMetPheIleMetThrAsnLysAla-----Ser 272  
Db 748 ATCAACCTTGACATCTCTTGTATCAACCGTACCTTGGCCATCGTGGGACCACTCT 807  
Qy 273 SerAsnGluProSerLysGlu-----LeuGluLysAspProValTrpLysLysLeu 290  
Db 808 AGCAACAACGGTGTCTAGAAAATGCCCTTATCGCTGAACACCT----- 852  
Qy 291 AsnAlaValLysAsnGlnArgValAspLysLeuAspArgAspLeuTrpAlaArgSerArg 310  
Db 853 ---GCTGTAATAATGTAAGATTATCAACCTTAACACCAAGACCTCTGTTGATCTAAGCGA 909  
Qy 311 GlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328  
Db 910 GCGGACTTGAATCAACAAAATCATGATTGAAGACATACAAAAAGCTTTGAAA 963

## RESULT 32

US-09-107-433-2206

; Sequence 2206, Application US/09107433

; Patent No. 6800744

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID

; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE

; THERAPEUTICS

; NUMBER OF SEQUENCES: 5206

; CORRESPONDENCE ADDRESS:

FOR DIAGN

ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: <Unknown>  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: <Unknown>  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,433  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085131  
FILING DATE: May 12, 1998  
APPLICATION NUMBER: 60/051553  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneka  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-011  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 2206:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 978 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pneumoniae  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (8) LOCATION 1...978  
SEQUENCE DESCRIPTION: SEQ ID NO: 2206:  
US-09-107-433-2206

## Alignment Scores:

Pred. No.:	4,696-19	Length:	978
Score:	250.00	Matches:	92
Percent Similarity:	46.45%	Conservative:	65
Best Local Similarity:	27.22%	Mismatches:	119
Query Match:	14.79%	Indels:	62
DB:	4	Gaps:	14

US-10-724-972A-6352 (1-335) x US-09-107-433-2206 (1-978)

Qy	20	LeuIleAlaThrAlaAlaCysGlyAsnAsnSerSerAsnSerSerLysGluSerSer	39
Db	61	TTGCTCTACTTGGTGCATGTAGTACAAACCTCAAGCACTAGTCAGACAGACAGACAGTAGC	120
Qy	40	LysAspGlyValGlu-----IleLysHisGluGluGlyThrThrLysValProLysHis	57
Db	121	TCTGCTCCACAGAGGTAACCATTAAGTTCACTGGACGAGGTCAAACTTTCCAAAGTT	180
Qy	58	ProLysArgValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspVal	77
Db	181	CCTGAAAAGATTGTGACCTTTCACCTCGCGCTCGGATACCTATTTCGCGCTTTAGGATTT	240
Qy	78	LysProValGlyIleAlaAspAsnLysLysAsnArgIleIleLysPro-----	94
Db	241	-----GAAAAAATATCGTGGAAATCGCTCGGAATGCCTACAAAAACT	273
Qy	95	-----LeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGln	110
Db	274	GTTCGAGCTTATCTAAAGACCTAGTGGAACTGTCAAAAATGTGTGTTTATGAAAGAA	333



Pred. No.: 1.11e-18 Length: 1119  
Score: 247.50 Matches: 95  
Percent Similarity: 46.76% Conservative: 71  
Best Local Similarity: 26.76% Mismatches: 130  
Query Match: 14.64% Indels: 59  
DB: 4 Gaps: 16

US-10-724-972A-6352 (1-335) x US-09-602-787A-597 (1-1119)

QY 9 Lysileu-----SerValileGlyLeuPheValleuileAlaThrAlaAaCys 26  
DB 119 AAAATCTGGCGAGTAGCGTGGAGGTTGCT---GTCTACCCCTGTTAGCTGGGTGT 175

QY 27 GlyAsnAsnSer-----SerAsnSerSerLysGluSerLysAspGly 42  
DB 176 TCTAACAAATGCAGATGACACCGAGTGTTCATCATCCACGCAAACTCCGCTTTTCT 235

QY 43 ValGluileLysHisGluGluGlyThrLysValProLysHisProLysArgValVal 62  
DB 236 GTTTCATTTGAACACGAGTTCCGAACACCAATCGATGATGATACCCGAAAGAGTTGTC 295

QY 63 ValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyile 82  
DB 296 ACCCTTGGCGTTACCGACCGCATATTGCTCGCATTTGGGACCGTCCAGTAGGCAAC 355

QY 83 AlaAspAsnLysLysAsnArgileileLysProLeuArgAspLysile-----Gly 100  
DB 356 ACCGGATACAAATCTCTCAAAACGAGTTGGGACCGTGCATGATGAGTTAGTGAAGGC 415

QY 101 Lys---TyrThrSerValGlyThrArgLysGlnProAsnLeuGluLysSerLysLeu 119  
DB 416 AAAGAAATTAACATGCTGCTGACTCTGATTCACACACAGATCTTGAACAAGTAGCAGCCCTG 475

QY 120 LysProAspLeuileileAlaAspAsnArgHisLysGly---IleTyrLysAspLeu 138  
DB 476 GAGCAGACCTGATTTAGGAGTCTCTCGGGGTTGACGAGCTGTGTATACGAGCACTA 535

QY 139 AsnLysIleAlaProThrileGluLeuLysSerPheAspGlyAspTyrAsn---GluAsn 157  
DB 536 TCTGATATCGCACCGGTGTCGCCCTCCAGCGGAAACAGTCGATACCGAGTAGCTGC 595

QY 158 IleAspAlaPheLysThrileSerLysAlaLeuGlyLysGluGluGlyLysLysArg 177  
DB 596 GAGGAAGCTACCAACTGTTGCCGTCGATGGGGCAATCAGAAAAAGCAACAGAGCTC 655

QY 178 LeuGluGluHisAspLysLysileGluGluTyrLysLysGluileThrMetAspLysAsn 197  
DB 656 AATGAGGAACAGATGCTGATCCAGCTGCGCGT-----GATGAAAAAT 700

QY 198 GlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisPro-----214  
DB 701 -----CCTTCTTTGACGGTAAACACGGAACCGTCATCTTGCCTACACAGGCT 748

QY 215 ---SerAsnSerTyrVal-----GlyGlnPheLeuSerGlnLeuGlyPhe 228  
DB 749 AAATACGGTGCTTACCTCCAGCGGATGACCGGGGCAATTCCTCGATTCATCTGGCAT 808

QY 229 -----LysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLysLys 245  
DB 809 TCGTGCCGGAAGCAGTTCTTTCGGAGACACCGCGCAGCTTC-----853

QY 246 GlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPhe 265  
DB 854 -----TTTGTGATGCTCCCGCTGAAAGCGTCAAGACGATAGCGGTGATGTTCTCCTC 907

QY 266 IleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLysAspPro 285  
DB 908 GTGCTT-----TCCAAACGACGAAATCTGGATATCACAGCAGAGAGATCCA 952

QY 286 ValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeu 305  
DB 953 CTGTTTGAACACATCAACGTTGTGCAAAAAAGACGCGAGTAATTTGTGGCAACACCGAA---1009

QY 306 TrpAlaArgSerArgGlyLeuile-----SerSer 315  
DB 1010 -----GAACGGGGGGGATTACTACAACCTAGTGTCTGTCTTCTTGTGCGTTG 1060

QY 316 GluGluMetAlaLysGluLeuValGluLeuSerLysLysAspSer 330  
DB 1061 GAACATCTCGCACCACTGATTGCTGAGGCTTTTGAAGTAAACTCA 1105

RESULT 34  
US-09-328-352-318  
; Sequence 318, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 318  
; LENGTH: 975  
; TYPE: DNA  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-318

Alignment Scores:  
Pred. No.: 5.05e-18 Length: 975  
Score: 241.00 Matches: 79  
Percent Similarity: 47.42% Conservative: 77  
Best Local Similarity: 24.01% Mismatches: 135  
Query Match: 14.26% Indels: 38  
DB: 4 Gaps: 10

US-10-724-972A-6352 (1-335) x US-09-328-352-318 (1-975)

QY 14 IleGlyLeuLeuPheValleuileAlaThrAlaAaCysGlyAsnAsnSerSerSerAsn 33  
DB 34 GTGCGCCCTCATCATAGCGCGAGCTCTCACTTACAGCATGTGATCAAAAGTTGCAGAT 93

QY 34 SerSerLysGluSerSerLys-----AspGlyValGluileLysHisGluGlyThr 51  
DB 94 ACCACTCAAGCTTCTCAGAGCTAGCAGACCGATTACGGTAAACATCGCGTTGGAAT 153

QY 52 ThrLysValProLysHisProLysArgValValLeuLeuGluTyrSerPheValAspAla 71  
DB 154 ACGGTGATAGATCATCTACCGCAACCGCTTGCCTGCTGATATGATGAATGAAGCCGACTTT 213

QY 72 LeuValAlaLeuAspValLysProValGlyileAlaAspAsnLysLysAsnArgile 91  
DB 214 CTAGATCAACTCAATGTCCTGATTTGGGAATGCCAAAGATTATGTTCCACACTTTTG 273

QY 92 IleLysProLeuArgAspLysileGlyLysTyrThrSerValGlyThrArgLysGlnPro 111  
DB 274 GAAAAATATAAAAGAT-----GCACAAATTCAGGATTTGGGTGCAATTGTACAGCCC 327

QY 112 AsnLeuGluGluileSerLysLysLeuLysProAspLeuileleAlaAspAsnAsnArgHis 131  
DB 328 AATATGGAAGGATATATGCAATTAACACACAGCTGATTTTGATG---ACGCCATTACAC 384

QY 132 LysGlyileTyrLysAspLeuAsnLysileAlaProThrileGluLeuLysSerPheAsp 151  
DB 385 GTTAATCAGTATCAGGAACCTGTCAAAATTTGCCCGCACCATTCAT-----TACGAT 435

QY 152 GlyAspTyrAsnGluAsn-----IleAspAlaPheLysThr 163  
DB 436 ATTAACCTTCAACATACGAGAGTAAATCATATTGCTGTGTTAAAGACCATATGATGACC 495

QY 164 IleSerLysAlaLeuGlyLysGluGluGlyLysArgLeuGluGluHisAspLys 183  
DB 496 TTAGGAAAAATATTATAAAGAGATTAGCCCGCCAGAAAGTTCTGAACTTGATGAA 555

QY 184 LysIleGluTyrLysLysGluileThrMetAspLysAsnGlnLysValLeuProAla 203

Alignment Scores: 6.36e-18 Length: 954  
Pred. No.: 240.00 Matches: 99  
Score: 240.00 Conservative: 64  
Percent Similarity: 45.28% Mismatches: 103  
Best Local Similarity: 27.50% Indels: 94  
Query Match: 14.20% Gaps: 20  
DB: 3

US-10-724-972A-6352 (1-335) x US-09-071-035-69 (1-954)

QY 8 LeuLysIleLeuSerValIleGly-----LeuLeuPheValLeuIleAlaThrAlaAa 25  
Db 25 TTAATAAAGAGCGTCTTAATTTGGTACACCCCTTCTTCTGGTTCATTCTTACTCGCAGCT 84  
QY 26 CysGlyAsn---AsnSerSerAsnSerSerLysGluSerSerLysAspGlyValGlu 44  
Db 85 TGTGTAATAACGTAATAAGAGCAACACGCTGACAAACA----- 126  
QY 45 IleLysHisGluGluGlyThr-----LysValProLysHisProLys 59  
Db 127 -----CATGAAGTAACAGATACCTTAGGCAATAAAGTAAACCGTCCCGGAAACCCAAA 180  
QY 60 ArgValValVal-----LeuGluTyrSerPheValAspAlaLeuValAlaLeuAspVal 77  
Db 181 CGGATTAATTCGGAGTTATTAGAA-----GATTATCTAGTTGCATTAGGAGAA 228  
QY 78 LysProValGly-----IleAlaAspAspAsnLysLysAsnArgIleIleLysPro 94  
Db 229 AAACCACTGGCACAATCGACAGTTGGCAAGCGACCATTCAGGATTATTAGCGAAAGAA 288  
QY 95 LeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsnLeu--- 113  
Db 289 TTGAAGAT-----GTCCCCCACTATTTC 312  
QY 114 -----GluGluIleSerLysLeuLysProAspLeuIleIleAlaAspAsn 128  
Db 313 TATGACTTGCATATGAAAGCGGTTCTTAAATTTGAACCTGACTTATTATTATCACTTCA 372  
QY 129 AsnArg-----HisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGlu 146  
Db 373 TCTGCTCTAGTTGAAGCGGTAATAACAAAGATACAGTAAATTTGCCCACTTATGTA 432  
QY 147 LeuLysSerPheAspGlyAspTyrAsnGluAsnIle-----AspAlaPheLysThr 163  
Db 433 GTCAAAAAC-----GGCGAAATGTCACTGGCGTGATCAATTTGGAAAGAT 477  
QY 164 IleSerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAsp--- 182  
Db 478 ATTGCCACTGTTTATAGATAAAGAGCAACGCAAAAAGTCTTAGAAGATTATCATACC 537  
QY 183 -----LysLysIleGluGluTyr-----LysLysGluIleThrMetAspLysAsnGln 198  
Db 538 TTAACCAAGGCGTCCAAAGATATCTTGGCAAAAAGAT----- 576  
QY 199 LysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyr 218  
Db 577 -----GCTGGCAATCTCGCGCAGTCTTATGGTAAACCAACCAACCAA 618  
QY 219 Val-----GlyGlnPheLeuSerGln---LeuGlyPhe 228  
Db 619 GTCTTTTATGGTTAGCGATATCGCTCAACGGAACCGTCTCTATCAGGACTTAGGCCTC 678  
QY 229 LysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLysLysGlyProTyr 248  
Db 679 CAA-----GTTCCAAAATTAGTGGAAAGAAATTTCTAAAAACCCCTACTCGCGGATTGG 729  
QY 249 LeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThr 268  
Db 730 AATCAAGTTCTTTTAGAAAAATTTAGCTGAGCTTGGCGAGACACACATTTTCTTGTAAC 769  
QY 269 AsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLysAspProValTyrLys 288  
Db 790 AGCGATGAATCAGCA-----CCTCTTTTCCAAAGCAATTTGGAAG 831

Db 556 CAAGTGAAGCAAGTA---CAAGCGGTAAACGGCCCAATCGCCAGAAAGAGCATTAGTCGTT 612  
QY 204 ValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeu 223  
Db 613 CTTTCAACAATCGGGCATTAGTAAATTTGGTATTTCAGTCAGCTATGCTTTATTTT 672  
QY 224 SerGlnLeuGlyPheLysGluAla-----LeuSerAspAspValThrLysGlyLeuSer 241  
Db 673 AATGCCTTTGGGTAAAGCCTGCAAGTGGTGGTGGATACAAAGCTACATGGG----- 726  
QY 242 LysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnPro 261  
Db 727 -----CAACCTATTTCGAGTGAATTTCATCAAAAAGCGGATCCC 765  
QY 262 GluArgMetPheIle-----MetThrAsnLysAlaSerSerAsnGlu 275  
Db 766 GATATCTTTATATCGTGACCGAAGTCTGTGATGAACATCGCTCAAAACATTAAATGCT 825  
QY 276 ProSerLeuLysGluLeuGluLysAspProValTyrLysLysLeuAsnAlaValLysAsn 295  
Db 826 GCCAGTGT-----GAAATCCCATTTATTCGTCAAAACCAAGCATGGAAAGAAC 873  
QY 296 GlnArgValAspIleLeuAspArgAspLeuTyrAlaArgSerArgGlyLeuIleSerSer 315  
Db 874 GGTGCTGTCAATTTTTCGATGCGGATCGCTGCTACACCGGCTGCAAGCCCAACCTCA 933  
QY 316 GluGluMetAlaLysGluLeuValGlu 324  
Db 934 CTCAAATAGTGAAGAGATGTGAAA 960

RESULT 35  
US-09-071-035-69  
; Sequence 69, Application US/09071035  
; Patent No. 6448043  
; GENERAL INFORMATION:  
; APPLICANT: Gil H. Choi  
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
; NUMBER OF SEQUENCES: 496  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/071,035  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: A. Anders Brooks  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB369P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 69:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 954 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-09-071-035-69

Qy 289 LysLeuAenAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArg 308  
Db 832 AACTTACTGCTGTGAAATAAACAAGTTTATACCTATGATAA-----AAA 879  
Qy 309 SerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerIysLys 328  
Db 880 AGTAGTTGTTATACACGACGACCTATTCCGAATACTCAATTTGTTGAAGATGTAATAAAA 939

## RESULT 36

US-09-902-540-5089  
; Sequence 5089, Application US/09902540  
; Patent No. 6833447

## GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B

; CURRENT APPLICATION NUMBER: US/09/902,540

; PRIOR FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/217,883

; NUMBER OF SEQ ID NOS: 2000-07-10

; SEQ ID NO 5089

; LENGTH: 837

; TYPE: DNA

; ORGANISM: Myxococcus xanthus

US-09-902-540-5089

## Alignment Scores:

Pred. No.: 8,87e-18 Length: 837  
Score: 238.00 Matches: 76  
Percent Similarity: 49.65% Conservative: 64  
Best Local Similarity: 26.95% Mismatches: 126  
Query Match: 14.08% Indels: 16  
DB: Gaps: 9

US-10-724-972A-6352 (1-335) x US-09-902-540-5089 (1-837)

Qy 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64  
Db 4 GTGACCGCAGGACGACGACCGCTGCTCCGCTGAAGCCGCGGGTGTGTGTTTC 63  
Qy 65 GluTySerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84  
Db 64 GACCTGTGTGCGCTGGACATCTCCAGGCCCTGGAGTGGAGCTCCAGCGGTGGCGGC 123  
Qy 85 AspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyThrSer 104  
Db 124 GACATGTTCCCGCCAGCACCTGACCGG---TTCCGGGAC-----GCCAAGTACCGCGC 174  
Qy 105 ValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIle 124  
Db 175 ATGGCACCTGTTTCGAGCCTGACTACGAGGCACCTCCAGCGCGGCGAAGCGGACCTCATC 234  
Qy 125 IleAlaAspAsnAsnArgHisLysGlyIleTyLysAspLeuAsnLysIleAlaProThr 144  
Db 235 ATCACC---GGGGCGCGCTCCAGCGCGAAGTACTCGAACCTCTCCGATCGCGCCACC 291  
Qy 145 IleGluLeuLysSerPheAspGlyAspTyThrAsnGluAsnIleAspAla---PheLysThr 163  
Db 292 ATCCAGCTGCGGATGAGCGGCAAGGACTTTCATCGCTCGGTGATTGCCAACACGAGATG 351  
Qy 164 IleSerLysAlaLeuGlyLysGluGluGluGlyLysArgLeuGluGluHisAspLys 183  
Db 352 CTACCCAGCTGTTTGGCAAGGAGGAGCGCGCGCTGATTGAGGACCTCGGCAAG 411  
Qy 184 LysIleGluGlyLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAla 203  
Db 412 TCGGTGGCGGACCTC---CAACAGATGACCGGACACCGGC---AAGGGGCTGTGCTGTG 465

Qy 204 ValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyThrValGlyGlnPheLeu 223  
Db 466 CTCGTACCGGAGCGCATGATGCGCTATGCGCGGTTTGGCGGTTCATCCAT 525  
Qy 224 SerGlnLeuGlyPheLysGluAlaLeuSerAspAspValThrLysGlyLeuSerLysTy 243  
Db 526 GGTGACTTGGCGTGGCCGAGCGGCTC-----GAGGGCTCGGACCTCG 570  
Qy 244 LeuLysGlyProTyThrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263  
Db 571 CTGACCGCGAG-----TCCATCAGCGCGAGTTTCATCGGGGAGAGAATCCCGACTGG 624  
Qy 264 MetPheIleMetThrAsnLysAlaSerSer-----AsnGluProSerLeuLysGluLeu 281  
Db 625 CTGTTCTGTCATCAGCGGACCGCGCCACCGGTGAGAAAGGCAATGCCCGCAGGTG 684  
Qy 282 GluLysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeu 301  
Db 685 CTGGACAATGAATTGGTGGCGGACGACGCGCTCGCAGAGAACAGGTCATCTACCTG 744  
Qy 302 AspArgAspLeuTrpAlaArgSerArgLysLeuIleSerSerGluGluMetAlaLysGlu 321  
Db 745 GACCCCGCGGTACCTACCTGACCGCGCGGCGGATCCAGTCCGTCAGCAGGCTCCCGGAC 804  
Qy 322 LeuVal 323  
Db 805 CAGGTC 810

## RESULT 37

US-08-961-083-23

; Sequence 23, Application US/08961083

; Patent No. 6159469

; GENERAL INFORMATION:

; APPLICANT: Choi et. al.

; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines

; NUMBER OF SEQUENCES: 452

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/961,083

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Brookes, A. Anders

; REGISTRATION NUMBER: 36,373

; REFERENCE/DOCKET NUMBER: PB340P2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 23:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 895 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

US-08-961-083-23

Alignment Scores:

Pred. No.: 1.66e-17 Length: 895

Score: 236.00 Matches: 88

Percent Similarity:	46.22%	Conservative:	65
Best Local Similarity:	26.59%	Mismatches:	112
Query Match:	13.96%	Indels:	66
DB:	3	Gaps:	14
US-10-724-972A-6352 (1-335) x US-08-961-083-23 (1-895)			
QY	29	AsnSerSerSerAsnSerSerLysGluSerSerLysAspGlyValGlu-----lleLys	46
DB	8	AACTCAAGCACTAGTCAGACAGACAGCAGTAGCTCTGCTCCACAGAGGTAAACATTAAA	67
QY	47	HisGluGluGlyThrThrLysValProLysHisProLysArgValValValLysGluTyr	66
DB	68	AGTTCACTCGACGAGGTCAAACTTCCAAAGTCTCTGAAAAGATTGTGACCTTTGACCTC	127
QY	67	SerPheValAspAlaLeuAlaLeuAspValLysProValGlyLleAlaAspAsn	86
DB	128	GGCGCTCGGATACTATTCGCGCTTTAGGATT-----	160
QY	87	LysLysAsnArgLlelleLysPro-----LeuArgAspLyslle	99
DB	161	GAATAAATAATATCGTCGAATGCTTACAAAACTGTTCCGACTTATCTAAAAGACCTAGTG	220
QY	100	GlyLysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluLysSerLysLeu	119
DB	221	GGAACTGTCAAAAATGTTGTTCTATGAAGAACCTGATTAGAGCTATCGCGCCCTT	280
QY	120	LysProAspLlellelleAlaAspAsnAsnArgHisGlyLleTyrLysAspLeuAsn	139
DB	281	GAGCTGATTGATTATTCGCTTCGCCAGTCACAAAAAATTCGTAGACAAA---TTCAAA	337
QY	140	LyslleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnleAsp	159
DB	338	GAATCGCCCCAACCGTTCTCTTCAAGCAAGCAGCAGCAGCTACTGGACTTCTACCAAG	397
QY	160	Ala---PheLysThrIleSerLysAlaLeuGlyLys-----GluGluGluGlyLysLys	176
DB	398	GCTAATATCGAATCTCTAGCAAGTCCTTCGCGGAACTGTTGACACAGAAAGCCAGAA	457
QY	177	ArgLeuGluGluHisAspLysLyslleGluGluTyr-----	188
DB	458	GAATTGACCAAGCTAGACAGAGCATCCAAAGAGTCGCTACTATAAATGAAGCTCTGAC	517
QY	189	LysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSer	208
DB	518	AAAAAAGCCCTTCGATCTCTCTTAATGAAGGAAAAATGGCAGCCCTTGGTCCCAATCT	577
QY	209	GlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLysSerGln---LeuGly	227
DB	578	CGTTTC-----TCTTCTTGTACCAAACTTGAAA	607
QY	228	PheLys-----GluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLys	245
DB	608	TTCAAAACCACTGATACAAAATTTGAGACTCAGCCACGGA-----	649
QY	246	GlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPhe	265
DB	650	-----CAAGAAGTCAGCTTTGAAAGTGTCAAAAGAAATCAACCCCTGACATCTCTTT	700
QY	266	IleMetThrAsnLysAla-----SerSerAsnGluProSerLeuLys	279
DB	701	GTCAATCAACCGTACCTTCGATCGTGGGACAACTCTAGCAACAGCGGTGCTTAGAA	760
QY	280	Glu-----LeuGluLysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArg	297
DB	761	AATGCCCTTATCGCTGAAACACT-----GCTGTAATAAATGGTAAG	802
QY	298	ValAspIleLeuAspArgAspLeuTrpAlaArgSerArgGlyLeulleSerSerGluGlu	317
DB	803	ATTATCAACTAACACCAGACCTCTGTTATCTAAGCGGAGCGGACTTGAATCAACAAA	862
QY	318	MetAlaLysGluLeuValGluLeuSerLysLys	328

DB

863

CTCATGATTGAAGACATACAAAAAGCTTTTGAAA

895

RESULT 38

US-09-536-784-23

Sequence 23, Application US/09536784

Patent No. 6573082

GENERAL INFORMATION:

APPLICANT: Choi et. al.

TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

NUMBER OF SEQUENCES: 452

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/536,784

FILING DATE: 30-Oct-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/961,083

FILING DATE: OCT-30-1997

ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks

REGISTRATION NUMBER: 41,971

REFERENCE/DOCKET NUMBER: PB340P3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 895 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 23:

US-09-536-784-23

Alignment Scores:

Pred. No.: 1,66e-17

Length: 895

Score: 236.00

Matches: 88

Percent Similarity: 46.22%

Conservative: 65

Best Local Similarity: 26.59%

Mismatches: 112

Query Match: 13.96%

Indels: 66

DB: 4

Gaps: 14

US-10-724-972A-6352 (1-335) x US-09-536-784-23 (1-895)

QY

29

AsnSerSerSerAsnSerSerLysGluSerSerLysAspGlyValGlu-----lleLys

46

DB

8

AACTCAAGCACTAGTCAGACAGACAGCAGTAGCTCTGCTCCACAGAGGTAAACATTAAA

67

QY

47

HisGluGluGlyThrThrLysValProLysHisProLysArgValValValLysGluTyr

66

DB

68

AGTTCACTCGACGAGGTCAAACTTCCAAAGTCTCTGAAAAGATTGTGACCTTTGACCTC

127

QY

67

SerPheValAspAlaLeuAlaLeuAspValLysProValGlyLleAlaAspAsn

86

DB

128

GGCGCTCGGATACTATTCGCGCTTTAGGATT-----

160

QY

87

LysLysAsnArgLlelleLysPro-----LeuArgAspLyslle

99

DB

161

GAATAAATAATATCGTCGAATGCTTACAAAACTGTTCCGACTTATCTAAAAGACCTAGTG

220

QY

100

GlyLysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluLysSerLysLeu

119



Db 221 GGAACGTGCAAAATGTTGGTTCTATGAAAGCACTGATTAGAAAGCTATATCGCCGCCCTT 280  
QY 120 LysProAspLeuIleAlaAspAsnArgHisLysGlyIleTyrLysAspLeuAsn 139  
Db 281 GAGCCTGATTGATTATCGCTTCGCCAGCTACACAAAATTCGTAGACAA--TTCAA 337  
QY 140 LysIleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAsp 159  
Db 338 GAAATCGCCCAACCGTCTCTTCCAGCAGCAAGGACGACTACTGGACTTCTTACCAAG 397  
QY 160 Ala---PheLysThrIleSerLysAlaLeuGlyLys-----GluGluGluGlyLysLys 176  
Db 398 GCTAATATCAATCTTAGCAAGTCTTCCGCGAAGCTGGTACACAGAAAGCCAGGAA 457  
QY 177 ArgLeuGluGluHisAspLysLysIleGluGluTyr----- 188  
Db 458 GAATTGACCAAGCTAGACAGAGCATCCAAGAAGTCGCTACTAAATGAAAGCTCTGCAC 517  
QY 189 LysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSer 208  
Db 518 AAAAAGCCCTTCGATCTCTTAAATGAAGGAAAAATGGCAGCCCTTGGTGCCAAATCT 577  
QY 209 GlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGln---LeuGly 227  
Db 578 CGTTTC-----TCYTTCTTGACCAAACTTGAAA 607  
QY 228 PheLys-----GluAlaLeuSerAspAspValThrLysGlyLeuSerLysTyrLeuLys 245  
Db 608 TTCAACCACTGATACAAAATTTGAAGACTCACGCCACGGA----- 649  
QY 246 GlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPhe 265  
Db 650 -----CAAGAGTCAGCTTTGAAAGTGTCAAAGAAATCAACCCCTGACATCTCTTT 700  
QY 266 IleMetThrAsnLysAla-----SerSerAsnGluProSerLeuLys 279  
Db 701 GTCATCAACCGTACCTTCCATCGTGGGCAACTCTAGCAACGACGGTCTCTAGAA 760  
QY 280 Glu-----LeuGluLysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArg 297  
Db 761 AATGCCCTTATCGCTGAACACCT-----GCTGCTAAATAATGGTAAG 802  
QY 298 ValAspIleLeuAspArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGlu 317  
Db 803 ATTATCCAACTAACACCGACCTCTGGTATCTTAAGCGGAGCGGACTTGAATCAACAAA 862  
QY 318 MetAlaLysGluLeuValGluLeuSerLysLys 328  
Db 863 CTCATGATTGAAGACATACAAAAGCTTTGAAA 895

## RESULT 39

US-09-134-000C-2609  
; Sequence 2609, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: LYNN Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US/09/134,000C  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2609  
; LENGTH: 906  
; TYPE: DNA  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-2609

Alignment Scores:

Pred. No.: 1.7e-17 Length: 906

Score: 236.00 Matches: 93  
Percent Similarity: 45.61% Conservative: 63  
Best Local Similarity: 27.19% Mismatches: 94  
Query Match: 13.96% Indels: 92  
DB: 4 Gaps: 19

US-10-724-972A-6352 (1-335) x US-09-134-000C-2609 (1-906)

QY 24 AlaAlaCysGlyAsn---AsnSerSerSerAsnSerSerLysGlySerSerLysAspGly 42  
Db 31 GCGCGCTTGTGTATACGAATAAAGAGCAACCAACCTGCACAAACA----- 78  
QY 43 ValGluIleLysHisGluGluGlyThr-----LysValProLysHis 57  
Db 79 -----CATGAAGTAAACAGATACCTTAGGCAATAAAGTAACCGTCCCGCGAAA 126  
QY 58 ProLysArgValValVal-----LeuGluTyrSerPheValAspAlaLeuValAlaLeu 75  
Db 127 CCCAAACCGATTATTTCGAGTTATTAGAA-----GATTATCTAGTTGCATTA 174  
QY 76 AspValLysProValGly-----IleAlaAspAsnLysLysAsnArgIleIle 92  
Db 175 GGAGAAAAACCAAGTGGCACAATGGACAGTTGGACAAGGCAGCATTCAGAGATTATTAGCG 234  
QY 93 LysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsn 112  
Db 235 AAGAAATTGAAAGAT-----GTCCCACT 258  
QY 113 Leu-----GluGluIleSerLysLeuLysProAspLeuIleIleAla 126  
Db 259 ATTCTCTATGACTTCCATATCAAGCGGTTCTAAATTTGAACCTGACTTATTATTAAATC 318  
QY 127 AspAsnAsnArg-----HisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144  
Db 319 AGTTTCATCTGCTCTAGTTGAAGCGGTAATAACAAGAATACAGATAAATTTGCGCCAACT 378  
QY 145 IleGluLeuLysSerPheAspLysAspTyrAsnGluAsnIle-----AspAlaPhe 161  
Db 379 TATGTAGTCAAAAAC-----GCCGAAAATGTCACTCGCGTGCATCAATTG 423  
QY 162 LysThrIleSerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluHis 181  
Db 424 GAAGATATTGCCACTGTTTATAGATANAAGAACCAAGCGAANAAGTGTTAGAAGATTAT 483  
QY 182 Asp-----LysLysIleGluGluTyr-----LysLysGluIleThrMetAspLys 196  
Db 484 GATACCTTAAACCAAGGCGTCCAGAATATCTTGCACAAAAGAT----- 528  
QY 197 AsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsn 216  
Db 529 -----GCTGGCAAAATCTGCGCAGTCTTATGGTAAACCAAC 564  
QY 217 SerTyrVal-----GlyGlnPheLeuSerGln---Leu 226  
Db 565 AACCAAGTCTTATGTTAGCGATAATCGCTCAAGCGGAACCGTGCTCTATCAGGACTTA 624  
QY 227 GlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGly 246  
Db 625 GGCCTCCAA-----GTTCCAAAATTAGTGAAGAATTTCTAAACCGCTACTCGC 675  
QY 247 ProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheIle 266  
Db 676 GATTGGAATCAAGTTTCTTTAGAAAAATTAGCTGAGCTTGACGACCAACATTTTCCT 735  
QY 267 MetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLysAspProVal 286  
Db 736 GTAAACAGCGATGAATCAGCA-----CCTCTTTTCCAAAGAACCAATT 777  
QY 287 TrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrp 306  
Db 778 TGAAGAAGAACTTACCTGCTGTGAAAAATAACCAAGTTCATACCTATGATAAA----- 828  
QY 307 AlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSer 326



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 10, 2005, 00:11:38 ; Search time 215 Seconds  
(without alignments)  
2549.547 Million cell updates/sec

Title: US-10-724-972A-6352  
Perfect score: 335  
Sequence: 1 GVESVRLKILSVIGLFLVL.....EEMAKELVLSKDKSKONK 335

Scoring table: OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2400006

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Command line parameters: -DEV=xlh  
-MODE=frame+pn.model -DEV=xlh  
-Q=/cgn2\_1/USPTO spool\_h/US10724972/runat\_03112005\_141928\_29427/app\_query.fasta\_1.519  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=oligo rn1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPTXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=1000 -DOICALIGN=200 -THR SCORE=quality -THR\_MIN=1 -ALIGN=40 -MODE=LOCAL  
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10724972@cgn 1 1 69 @runat\_03112005\_141928\_29427 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	51	15.2	183	3	US-09-134-001C-2023
2	35	10.4	3046	4	Sequence 2023, Ap
3	26	7.8	242	4	Sequence 3782, Ap
4	26	7.8	242	4	Sequence 2556, Ap
5	23	6.9	668	4	Sequence 2556, Ap
6	23	6.9	668	4	Sequence 355, App
7	14	4.2	999	4	Sequence 15, App
8	14	4.2	999	4	Sequence 15, Appl
9	14	4.2	3775	4	Sequence 15, Appl
10	14	4.2	3775	4	Sequence 238, App
11	9	2.7	13248	4	Sequence 238, App
12	9	2.7	13248	4	Sequence 12024, A
					Sequence 14230, A

C 13	9	2.7	203475	4	US-09-949-016-14516	Sequence 14516, A
C 14	9	2.7	203475	4	US-09-949-016-14517	Sequence 14517, A
C 15	9	2.7	203475	4	US-09-949-016-14518	Sequence 14518, A
C 16	9	2.7	203475	4	US-09-949-016-14519	Sequence 14519, A
C 17	9	2.7	203475	4	US-09-949-016-17226	Sequence 17226, A
C 18	9	2.7	203475	4	US-09-949-016-17227	Sequence 17227, A
C 19	9	2.7	203475	4	US-09-949-016-17228	Sequence 17228, A
C 20	9	2.7	203475	4	US-09-949-016-17229	Sequence 17229, A
C 21	8	2.4	181	4	US-09-513-999C-3507	Sequence 3507, A
C 22	8	2.4	305	4	US-09-513-999C-3507	Sequence 3507, A
C 23	8	2.4	352	4	US-09-513-999C-3507	Sequence 3507, A
C 24	8	2.4	378	4	US-09-513-999C-3507	Sequence 3507, A
C 25	8	2.4	390	4	US-09-513-999C-3507	Sequence 3507, A
C 26	8	2.4	564	4	US-09-513-999C-3507	Sequence 3507, A
C 27	8	2.4	601	4	US-09-513-999C-3507	Sequence 3507, A
C 28	8	2.4	601	4	US-09-513-999C-3507	Sequence 3507, A
C 29	8	2.4	601	4	US-09-513-999C-3507	Sequence 3507, A
C 30	8	2.4	601	4	US-09-513-999C-3507	Sequence 3507, A
C 31	8	2.4	601	4	US-09-513-999C-3507	Sequence 3507, A
C 32	8	2.4	789	3	US-09-134-001C-909	Sequence 909, App
C 33	8	2.4	789	3	US-09-134-001C-909	Sequence 909, App
C 34	8	2.4	843	4	US-09-543-681A-1850	Sequence 1850, App
C 35	8	2.4	843	4	US-09-543-681A-1850	Sequence 1850, App
C 36	8	2.4	886	4	US-09-270-767-14140	Sequence 14140, A
C 37	8	2.4	981	4	US-09-489-039A-2246	Sequence 2246, App
C 38	8	2.4	1001	3	US-09-641-638-192	Sequence 192, App
C 39	8	2.4	1001	3	US-09-641-638-192	Sequence 192, App
C 40	8	2.4	1001	4	US-10-170-097-193	Sequence 193, App
C 41	8	2.4	1389	4	US-09-489-039A-3577	Sequence 3577, App
C 42	8	2.4	1470	4	US-09-902-540-133	Sequence 133, App
C 43	8	2.4	1491	4	US-09-134-000C-497	Sequence 497, App
C 44	8	2.4	1875	3	US-09-422-869-21	Sequence 21, Appl
C 45	8	2.4	2064	4	US-09-620-312D-375	Sequence 375, App
C 46	8	2.4	2103	4	US-09-620-312D-375	Sequence 375, App
C 47	8	2.4	2148	4	US-09-620-312D-374	Sequence 374, App
C 48	8	2.4	2283	4	US-09-248-796A-3465	Sequence 3465, App
C 49	8	2.4	3047	4	US-09-710-279-4113	Sequence 4113, App
C 50	8	2.4	3100	4	US-09-270-767-12855	Sequence 12855, A
C 51	8	2.4	3572	4	US-09-270-767-11151	Sequence 11151, A
C 52	8	2.4	3821	4	US-08-956-171E-304	Sequence 304, App
C 53	8	2.4	3821	4	US-08-956-171E-304	Sequence 304, App
C 54	8	2.4	5648	3	US-09-371-008-1	Sequence 1, Appl
C 55	8	2.4	13015	4	US-09-639-207-10	Sequence 10, Appl
C 56	8	2.4	17607	4	US-09-949-016-15968	Sequence 15968, A
C 57	8	2.4	35524	3	US-08-923-137-1	Sequence 1, Appl
C 58	8	2.4	45136	3	US-09-422-869-1	Sequence 1, Appl
C 59	8	2.4	54245	4	US-09-949-016-13499	Sequence 13499, A
C 60	8	2.4	134987	4	US-09-949-016-15348	Sequence 15348, A
C 61	8	2.4	134987	4	US-09-949-016-15349	Sequence 15349, A
C 62	8	2.4	134987	4	US-09-949-016-15350	Sequence 15350, A
C 63	8	2.4	134987	4	US-09-949-016-15507	Sequence 15507, A
C 64	8	2.4	134987	4	US-09-949-016-15508	Sequence 15508, A
C 65	8	2.4	134987	4	US-09-949-016-15509	Sequence 15509, A
C 66	8	2.4	146307	4	US-09-949-016-14881	Sequence 14881, A
C 67	8	2.4	146307	4	US-09-949-016-14882	Sequence 14882, A
C 68	8	2.4	146307	4	US-09-949-016-14883	Sequence 14883, A
C 69	8	2.4	146307	4	US-09-949-016-14884	Sequence 14884, A
C 70	8	2.4	146307	4	US-09-949-016-14885	Sequence 14885, A
C 71	8	2.4	146307	4	US-09-949-016-14886	Sequence 14886, A
C 72	8	2.4	146307	4	US-09-949-016-14887	Sequence 14887, A
C 73	8	2.4	146307	4	US-09-949-016-14888	Sequence 14888, A
C 74	8	2.4	184805	4	US-09-949-016-11747	Sequence 11747, A
C 75	8	2.4	184805	4	US-09-949-016-12835	Sequence 12835, A
C 76	8	2.4	184805	4	US-09-949-016-12836	Sequence 12836, A
C 77	8	2.4	184805	4	US-09-949-016-12837	Sequence 12837, A
C 78	8	2.4	175236	4	US-09-949-016-16513	Sequence 16513, A
C 79	8	2.4	177293	4	US-09-949-016-16513	Sequence 16513, A
C 80	8	2.4	187169	4	US-09-949-016-12776	Sequence 12776, A
C 81	8	2.4	191569	4	US-09-949-016-15940	Sequence 15940, A
C 82	8	2.4	256171	4	US-09-949-016-12822	Sequence 12822, A
C 83	8	2.4	256176	4	US-09-949-016-15524	Sequence 15524, A
C 84	8	2.4	4403765	3	US-09-103-840A-2	Sequence 2, Appl
C 85	8	2.4	4403765	3	US-09-103-840A-2	Sequence 2, Appl

86	8	2.4	4411529	3	US-09-103-840A-1	Sequence 1, Appli	7	2.1	498	4	US-09-252-991A-15197	Sequence 15197, A
C 87	8	2.4	4411529	3	US-09-103-840A-1	Sequence 1, Appli	7	2.1	503	4	US-09-621-976-17139	Sequence 17139, A
C 88	7	2.1	21	4	US-09-422-978-11680	Sequence 11680, A	7	2.1	504	4	US-09-248-796A-5217	Sequence 5217, Ap
C 89	7	2.1	25	4	US-09-396-196G-64569	Sequence 64569, A	7	2.1	506	4	US-09-513-999C-9805	Sequence 9805, Ap
C 90	7	2.1	97	5	PCT-US94-06456-26	Sequence 26, Appl	7	2.1	514	4	US-09-270-767-7141	Sequence 7141, Ap
C 91	7	2.1	97	5	PCT-US94-06456-55	Sequence 55, Appl	7	2.1	514	4	US-09-270-767-7141	Sequence 7141, Ap
C 92	7	2.1	147	4	US-09-328-352-3227	Sequence 3227, Ap	7	2.1	533	6	5482709-5	Patent No. 5482709
C 93	7	2.1	183	4	US-09-248-796A-12518	Sequence 12518, A	7	2.1	533	6	5482709-5	Patent No. 5482709
C 94	7	2.1	192	4	US-09-513-999C-16968	Sequence 16968, A	7	2.1	543	6	5273901-6	Patent No. 5273901
C 95	7	2.1	194	4	US-09-513-999C-16968	Sequence 16968, A	7	2.1	543	6	5273901-6	Patent No. 5273901
C 96	7	2.1	207	4	US-09-248-796A-12486	Sequence 12486, A	7	2.1	553	3	US-09-385-982-431	Sequence 431, App
C 97	7	2.1	207	4	US-09-248-796A-9966	Sequence 9966, Ap	7	2.1	565	3	US-09-710-279-2657	Sequence 2697, Ap
C 98	7	2.1	210	4	US-09-248-796A-9966	Sequence 9966, Ap	7	2.1	567	4	US-09-248-796A-6399	Sequence 6399, Ap
C 99	7	2.1	220	4	US-09-248-796A-9966	Sequence 9966, Ap	7	2.1	573	4	US-09-134-001C-2549	Sequence 2549, Ap
C 100	7	2.1	231	4	US-09-107-433-440	Sequence 433, Appl	7	2.1	573	3	US-09-248-796A-13766	Sequence 13766, A
C 101	7	2.1	234	4	US-09-621-976-410	Sequence 410, App	7	2.1	579	3	US-09-248-796A-13766	Sequence 13766, A
C 102	7	2.1	237	4	US-09-621-976-410	Sequence 410, App	7	2.1	582	4	US-09-252-991A-8856	Sequence 8856, Ap
C 103	7	2.1	239	4	US-09-621-976-13790	Sequence 13790, A	7	2.1	584	4	US-09-543-681A-2177	Sequence 2177, Ap
C 104	7	2.1	240	4	US-09-489-039A-2474	Sequence 2474, Ap	7	2.1	594	4	US-09-543-681A-2177	Sequence 2177, Ap
C 105	7	2.1	243	3	US-09-134-001C-908	Sequence 908, App	7	2.1	597	4	US-09-598-401C-100	Sequence 100, App
C 106	7	2.1	272	4	US-09-513-999C-33598	Sequence 33598, A	7	2.1	600	4	US-09-408-020-21	Sequence 21, Appl
C 107	7	2.1	277	4	US-09-313-294A-4117	Sequence 4117, Ap	7	2.1	601	4	US-09-949-016-19666	Sequence 19666, A
C 108	7	2.1	282	4	US-09-489-039A-4520	Sequence 4520, Ap	7	2.1	601	4	US-09-949-016-21191	Sequence 21191, A
C 109	7	2.1	292	4	US-09-513-999C-27614	Sequence 27614, A	7	2.1	601	4	US-09-949-016-27815	Sequence 27815, A
C 110	7	2.1	294	4	US-09-543-681A-2569	Sequence 2569, Ap	7	2.1	601	4	US-09-949-016-32695	Sequence 32695, A
C 111	7	2.1	306	4	US-09-252-991A-14703	Sequence 14703, A	7	2.1	601	4	US-09-949-016-32942	Sequence 32942, A
C 112	7	2.1	306	4	US-09-328-352-2197	Sequence 2197, Ap	7	2.1	601	4	US-09-949-016-35342	Sequence 35342, A
C 113	7	2.1	312	4	US-09-489-039A-3085	Sequence 3085, Ap	7	2.1	601	4	US-09-949-016-47058	Sequence 47058, A
C 114	7	2.1	324	1	US-08-592-126-120	Sequence 126, App	7	2.1	601	4	US-09-949-016-53986	Sequence 53986, A
C 115	7	2.1	324	4	US-09-168-595-120	Sequence 120, App	7	2.1	601	4	US-09-949-016-55127	Sequence 55127, A
C 116	7	2.1	325	4	US-09-513-999C-35228	Sequence 35228, A	7	2.1	601	4	US-09-949-016-65234	Sequence 65234, A
C 117	7	2.1	328	4	US-09-471-276-254	Sequence 254, App	7	2.1	601	4	US-09-949-016-69685	Sequence 69685, A
C 118	7	2.1	332	4	US-09-513-999C-1578	Sequence 1578, Ap	7	2.1	601	4	US-09-949-016-69686	Sequence 69686, A
C 119	7	2.1	336	4	US-09-902-540-3936	Sequence 3936, Ap	7	2.1	601	4	US-09-949-016-70408	Sequence 70408, A
C 120	7	2.1	373	3	US-09-280-116-25	Sequence 5, Appli	7	2.1	601	4	US-09-949-016-73301	Sequence 73301, A
C 121	7	2.1	379	1	US-08-145-617-5	Sequence 5, Appli	7	2.1	601	4	US-09-949-016-74851	Sequence 74851, A
C 122	7	2.1	384	4	US-09-902-540-7527	Sequence 7527, Ap	7	2.1	601	4	US-09-949-016-75050	Sequence 75050, A
C 123	7	2.1	387	4	US-09-513-999C-2236	Sequence 2236, Ap	7	2.1	601	4	US-09-949-016-75249	Sequence 75249, A
C 124	7	2.1	399	4	US-09-583-110-1470	Sequence 1470, Ap	7	2.1	601	4	US-09-949-016-75448	Sequence 75448, A
C 125	7	2.1	400	4	US-09-355-709C-6	Sequence 6, Appli	7	2.1	601	4	US-09-949-016-75647	Sequence 75647, A
C 126	7	2.1	400	4	US-08-956-171E-3743	Sequence 3743, Ap	7	2.1	601	4	US-09-949-016-76149	Sequence 76149, A
C 127	7	2.1	400	4	US-08-956-171E-4051	Sequence 4051, Ap	7	2.1	601	4	US-09-949-016-81554	Sequence 81554, A
C 128	7	2.1	400	4	US-08-781-986A-3743	Sequence 3743, Ap	7	2.1	601	4	US-09-949-016-81555	Sequence 81555, A
C 129	7	2.1	400	4	US-08-781-986A-4051	Sequence 4051, Ap	7	2.1	601	4	US-09-949-016-81557	Sequence 81557, A
C 130	7	2.1	414	4	US-09-252-991A-8746	Sequence 8746, Ap	7	2.1	601	4	US-09-949-016-84956	Sequence 84956, A
C 131	7	2.1	414	4	US-09-107-433-1001	Sequence 1001, Ap	7	2.1	601	4	US-09-949-016-86791	Sequence 86791, A
C 132	7	2.1	422	4	US-09-621-976-8754	Sequence 8754, Ap	7	2.1	601	4	US-09-949-016-95315	Sequence 95315, A
C 133	7	2.1	423	1	US-08-470-179-66	Sequence 66, Appl	7	2.1	601	4	US-09-949-016-95493	Sequence 95493, A
C 134	7	2.1	423	1	US-08-470-179-66	Sequence 66, Appl	7	2.1	601	4	US-09-949-016-95671	Sequence 95671, A
C 135	7	2.1	432	3	US-09-397-787-207	Sequence 207, App	7	2.1	601	4	US-09-949-016-95849	Sequence 95849, A
C 136	7	2.1	432	3	US-09-252-991A-3546	Sequence 3546, Ap	7	2.1	601	4	US-09-949-016-106205	Sequence 106205, A
C 137	7	2.1	435	4	US-09-248-796A-8821	Sequence 8821, Ap	7	2.1	601	4	US-09-949-016-106404	Sequence 106404, A
C 138	7	2.1	440	4	US-09-270-767-3111	Sequence 3111, Ap	7	2.1	601	4	US-09-949-016-106603	Sequence 106603, A
C 139	7	2.1	440	4	US-09-270-767-18593	Sequence 18593, A	7	2.1	601	4	US-09-949-016-106802	Sequence 106802, A
C 140	7	2.1	441	4	US-09-513-999C-648	Sequence 648, App	7	2.1	601	4	US-09-949-016-107001	Sequence 107001, A
C 141	7	2.1	462	3	US-09-643-597-253	Sequence 253, App	7	2.1	601	4	US-09-949-016-107043	Sequence 107043, A
C 142	7	2.1	462	3	US-09-480-884A-253	Sequence 253, App	7	2.1	601	4	US-09-949-016-107085	Sequence 107085, A
C 143	7	2.1	462	4	US-09-542-615A-253	Sequence 253, App	7	2.1	601	4	US-09-949-016-107127	Sequence 107127, A
C 144	7	2.1	462	4	US-09-606-421B-253	Sequence 253, App	7	2.1	601	4	US-09-949-016-107169	Sequence 107169, A
C 145	7	2.1	462	4	US-09-476-496A-253	Sequence 253, App	7	2.1	601	4	US-09-949-016-107211	Sequence 107211, A
C 146	7	2.1	462	4	US-09-476-496A-253	Sequence 253, App	7	2.1	601	4	US-09-949-016-107253	Sequence 107253, A
C 147	7	2.1	464	4	US-09-630-940B-253	Sequence 253, App	7	2.1	601	4	US-09-949-016-126784	Sequence 126784, A
C 148	7	2.1	464	4	US-09-621-976-14241	Sequence 14241, A	7	2.1	601	4	US-09-949-016-126833	Sequence 126833, A
C 149	7	2.1	465	4	US-09-252-991A-8912	Sequence 8912, Ap	7	2.1	601	4	US-09-949-016-136882	Sequence 136882, A
C 150	7	2.1	482	4	US-09-702-705-843	Sequence 843, App	7	2.1	601	4	US-09-949-016-134417	Sequence 134417, A
C 151	7	2.1	482	4	US-09-736-457-843	Sequence 843, App	7	2.1	601	4	US-09-949-016-134466	Sequence 134466, A
C 152	7	2.1	482	4	US-09-614-124B-843	Sequence 843, App	7	2.1	601	4	US-09-949-016-134515	Sequence 134515, A
C 153	7	2.1	482	4	US-09-671-325-843	Sequence 843, App	7	2.1	601	4	US-09-949-016-135704	Sequence 135704, A
C 154	7	2.1	482	4	US-09-658-824-843	Sequence 843, App	7	2.1	601	4	US-09-949-016-142415	Sequence 142415, A
C 155	7	2.1	486	4	US-09-270-767-2618	Sequence 2618, Ap	7	2.1	601	4	US-09-949-016-144728	Sequence 144728, A
C 156	7	2.1	486	4	US-09-270-767-17900	Sequence 17900, A	7	2.1	601	4	US-09-949-016-147691	Sequence 147691, A
C 157	7	2.1	486	4	US-09-639-207-13	Sequence 13, Appl	7	2.1	601	4	US-09-949-016-147691	Sequence 147691, A
C 158	7	2.1	489	4	US-09-248-796A-7305	Sequence 7305, Ap	7	2.1	601	4	US-09-949-016-150281	Sequence 150281, A

C 232	7	2.1	601	4	US-09-949-016-150328	Sequence 150328,	305	7	2.1	935	4	US-09-513-999C-967	Sequence 967, App
C 233	7	2.1	601	4	US-09-949-016-153691	Sequence 153691,	306	7	2.1	942	4	US-09-902-540-2655	Sequence 2655, App
C 234	7	2.1	601	4	US-09-949-016-155308	Sequence 155308,	307	7	2.1	948	4	US-09-452-991A-9121	Sequence 9121, App
C 235	7	2.1	601	4	US-09-949-016-155334	Sequence 155334,	308	7	2.1	951	4	US-09-543-681A-1701	Sequence 1701, App
C 236	7	2.1	601	4	US-09-949-016-158632	Sequence 158632,	309	7	2.1	963	4	US-09-710-279-1785	Sequence 1785, App
C 237	7	2.1	601	4	US-09-949-016-180601	Sequence 180601,	310	7	2.1	966	4	US-09-252-991A-15581	Sequence 15581, A
C 238	7	2.1	601	4	US-09-949-016-180602	Sequence 180602,	311	7	2.1	975	3	US-09-071-035-37	Sequence 37, Appl
C 239	7	2.1	601	4	US-09-949-016-160602	Sequence 160602,	312	7	2.1	978	4	US-09-543-681A-1393	Sequence 1393, App
C 240	7	2.1	601	4	US-09-949-016-162287	Sequence 162287,	313	7	2.1	984	4	US-09-248-796A-5978	Sequence 5978, App
C 241	7	2.1	601	4	US-09-949-016-166486	Sequence 166486,	314	7	2.1	986	4	US-09-134-000C-937	Sequence 937, App
C 242	7	2.1	601	4	US-09-949-016-166487	Sequence 166487,	315	7	2.1	986	4	US-09-270-767-2302	Sequence 2302, App
C 243	7	2.1	601	4	US-09-949-016-168100	Sequence 168100,	316	7	2.1	986	4	US-09-270-767-17584	Sequence 17584, A
C 244	7	2.1	601	4	US-09-949-016-169415	Sequence 169415,	317	7	2.1	993	3	US-09-134-001C-1788	Sequence 1788, App
C 245	7	2.1	601	4	US-09-949-016-169416	Sequence 169416,	318	7	2.1	998	3	US-08-858-207A-150	Sequence 150, App
C 246	7	2.1	601	4	US-09-949-016-169417	Sequence 169417,	319	7	2.1	1008	4	US-09-248-796A-2396	Sequence 2396, App
C 247	7	2.1	601	4	US-09-949-016-170440	Sequence 170440,	320	7	2.1	1024	4	US-09-328-475C-85	Sequence 85, Appl
C 248	7	2.1	601	4	US-09-949-016-170441	Sequence 170441,	321	7	2.1	1038	4	US-09-248-796A-456	Sequence 456, App
C 249	7	2.1	601	4	US-09-949-016-177593	Sequence 177593,	322	7	2.1	1038	4	US-10-121-857-70	Sequence 70, Appl
C 250	7	2.1	601	4	US-09-949-016-177594	Sequence 177594,	323	7	2.1	1053	4	US-09-270-767-11001	Sequence 2261, App
C 251	7	2.1	601	4	US-09-949-016-177597	Sequence 177597,	324	7	2.1	1072	4	US-09-489-039A-2261	Sequence 2261, App
C 252	7	2.1	601	4	US-09-949-016-179109	Sequence 179109,	325	7	2.1	1089	4	US-09-270-767-10841	Sequence 10841, A
C 253	7	2.1	601	4	US-09-949-016-179110	Sequence 179110,	326	7	2.1	1089	4	US-09-248-796A-1700	Sequence 1700, App
C 254	7	2.1	601	4	US-09-949-016-180642	Sequence 180642,	327	7	2.1	1095	4	US-09-543-681A-4027	Sequence 4027, App
C 255	7	2.1	601	4	US-09-949-016-185772	Sequence 185772,	328	7	2.1	1124	4	US-09-252-991A-8798	Sequence 8798, App
C 256	7	2.1	601	4	US-09-949-016-186879	Sequence 186879,	329	7	2.1	1131	4	US-09-107-532A-2794	Sequence 2794, App
C 257	7	2.1	601	4	US-09-949-016-190164	Sequence 190164,	330	7	2.1	1135	4	US-09-203-958A-3	Sequence 3, Appl
C 258	7	2.1	601	4	US-09-949-016-190165	Sequence 190165,	331	7	2.1	1143	4	US-09-270-767-12606	Sequence 12606, A
C 259	7	2.1	601	4	US-09-949-016-191								

C 378	7	2.1	1563	3	US-09-292-858B-11	Sequence 11, Appl	451	7	2.1	1959	3	US-08-996-441B-43	Sequence 43, Appl
C 379	7	2.1	1584	4	US-08-107-532A-1168	Sequence 1168, Ap	452	7	2.1	1959	3	US-08-996-441B-45	Sequence 45, Appl
C 380	7	2.1	1584	4	US-09-107-532A-124	Sequence 124, Ap	453	7	2.1	1959	3	US-08-996-441B-47	Sequence 47, Appl
C 381	7	2.1	1585	4	US-09-023-655-1329	Sequence 1329, Ap	454	7	2.1	1959	3	US-08-996-441B-49	Sequence 49, Appl
C 382	7	2.1	1632	4	US-09-252-991A-15621	Sequence 15621, A	455	7	2.1	1959	3	US-08-996-441B-53	Sequence 53, Appl
C 383	7	2.1	1641	4	US-09-489-039A-2456	Sequence 2456, Ap	456	7	2.1	1959	3	US-08-996-441B-59	Sequence 59, Appl
C 384	7	2.1	1655	4	US-09-248-796A-6368	Sequence 6368, Ap	457	7	2.1	1959	3	US-08-996-441B-61	Sequence 61, Appl
C 385	7	2.1	1683	4	US-09-552-991A-9428	Sequence 9428, Ap	458	7	2.1	1959	3	US-08-996-441B-63	Sequence 63, Appl
C 386	7	2.1	1689	4	US-09-902-540-2528	Sequence 2528, Ap	459	7	2.1	1959	3	US-08-996-441B-65	Sequence 65, Appl
C 387	7	2.1	1719	3	US-09-160-065-1	Sequence 1, Appl1	460	7	2.1	1959	3	US-08-996-441B-67	Sequence 67, Appl
C 388	7	2.1	1722	4	US-09-902-540-3804	Sequence 3804, Ap	461	7	2.1	1959	3	US-08-996-441B-97	Sequence 97, Appl
C 389	7	2.1	1770	4	US-08-976-063B-29	Sequence 29, Appl	462	7	2.1	1959	3	US-08-993-722A-1	Sequence 1, Appl1
C 390	7	2.1	1771	4	US-09-252-991A-9460	Sequence 9460, Ap	463	7	2.1	1959	3	US-08-993-722A-3	Sequence 3, Appl1
C 391	7	2.1	1824	4	US-09-540-236-86	Sequence 86, Appl	464	7	2.1	1959	3	US-08-993-722A-5	Sequence 5, Appl1
C 392	7	2.1	1830	4	US-09-252-991A-3563	Sequence 3563, Ap	465	7	2.1	1959	3	US-08-993-722A-7	Sequence 7, Appl1
C 393	7	2.1	1838	3	US-09-227-357-32	Sequence 3, Appl	466	7	2.1	1959	3	US-08-993-722A-9	Sequence 9, Appl1
C 394	7	2.1	1843	3	US-08-718-738-1	Sequence 1, Appl1	467	7	2.1	1959	3	US-08-993-722A-11	Sequence 11, Appl
C 395	7	2.1	1843	3	US-09-221-844-1	Sequence 1, Appl1	468	7	2.1	1959	3	US-08-993-722A-13	Sequence 13, Appl
C 396	7	2.1	1843	4	US-09-843-846-1	Sequence 1, Appl1	469	7	2.1	1959	3	US-08-993-722A-15	Sequence 15, Appl
C 397	7	2.1	1843	5	FCT-US95-0332A-1	Sequence 11207, A	470	7	2.1	1959	3	US-08-993-722A-17	Sequence 17, Appl
C 398	7	2.1	1845	4	US-09-252-991A-11207	Sequence 467, App	471	7	2.1	1959	3	US-08-993-722A-19	Sequence 19, Appl
C 399	7	2.1	1864	3	US-09-071-035-467	Sequence 11062, A	472	7	2.1	1959	3	US-08-993-722A-21	Sequence 21, Appl
C 400	7	2.1	1891	4	US-09-270-767-11062	Sequence 1429, Ap	473	7	2.1	1959	3	US-08-993-722A-23	Sequence 23, Appl
C 401	7	2.1	1896	4	US-09-252-991A-1429	Sequence 50, Appl	474	7	2.1	1959	3	US-08-993-722A-25	Sequence 25, Appl
C 402	7	2.1	1908	3	US-08-845-258-50	Sequence 50, Appl	475	7	2.1	1959	3	US-08-993-722A-27	Sequence 27, Appl
C 403	7	2.1	1908	3	US-08-990-571-50	Sequence 50, Appl	476	7	2.1	1959	3	US-08-993-722A-29	Sequence 29, Appl
C 404	7	2.1	1908	3	US-09-528-784A-50	Sequence 50, Appl	477	7	2.1	1959	3	US-08-993-722A-31	Sequence 31, Appl
C 405	7	2.1	1908	4	US-09-569-098A-50	Sequence 1, Appl1	478	7	2.1	1959	3	US-08-993-722A-33	Sequence 33, Appl
C 406	7	2.1	1920	1	US-08-186-222-1	Sequence 1, Appl1	479	7	2.1	1959	3	US-08-993-722A-35	Sequence 35, Appl
C 407	7	2.1	1926	4	US-09-252-991A-3531	Sequence 3531, Ap	480	7	2.1	1959	3	US-08-993-722A-37	Sequence 37, Appl
C 408	7	2.1	1938	4	US-09-949-016-5558	Sequence 5558, Ap	481	7	2.1	1959	3	US-08-993-722A-39	Sequence 39, Appl
C 409	7	2.1	1953	1	US-08-315-468-5	Sequence 5, Appl1	482	7	2.1	1959	3	US-08-993-722A-41	Sequence 41, Appl
C 410	7	2.1	1953	3	US-07-941-650A-3	Sequence 3, Appl1	483	7	2.1	1959	3	US-08-993-722A-43	Sequence 43, Appl
C 411	7	2.1	1955	4	US-09-949-016-770	Sequence 770, App	484	7	2.1	1959	3	US-08-993-722A-45	Sequence 45, Appl
C 412	7	2.1	1956	3	US-08-996-441B-51	Sequence 51, Appl	485	7	2.1	1959	3	US-08-993-722A-47	Sequence 47, Appl
C 413	7	2.1	1956	3	US-08-996-441B-55	Sequence 55, Appl	486	7	2.1	1959	3	US-08-993-722A-49	Sequence 49, Appl
C 414	7	2.1	1956	3	US-08-996-441B-57	Sequence 57, Appl	487	7	2.1	1959	3	US-08-993-722A-51	Sequence 51, Appl
C 415	7	2.1	1956	3	US-08-993-722A-51	Sequence 51, Appl	488	7	2.1	1959	3	US-08-993-722A-53	Sequence 53, Appl
C 416	7	2.1	1956	3	US-08-993-722A-55	Sequence 55, Appl	489	7	2.1	1959	3	US-08-993-722A-55	Sequence 55, Appl
C 417	7	2.1	1956	3	US-08-993-722A-57	Sequence 57, Appl	490	7	2.1	1959	3	US-08-993-722A-59	Sequence 59, Appl
C 418	7	2.1	1956	3	US-08-993-170A-51	Sequence 51, Appl	491	7	2.1	1959	3	US-08-993-722A-61	Sequence 61, Appl
C 419	7	2.1	1956	3	US-08-993-170A-55	Sequence 55, Appl	492	7	2.1	1959	3	US-08-993-722A-63	Sequence 63, Appl
C 420	7	2.1	1956	3	US-08-993-170A-57	Sequence 57, Appl	493	7	2.1	1959	3	US-08-993-722A-65	Sequence 65, Appl
C 421	7	2.1	1956	3	US-08-993-775B-51	Sequence 51, Appl	494	7	2.1	1959	3	US-08-993-722A-67	Sequence 67, Appl
C 422	7	2.1	1956	3	US-08-993-775B-55	Sequence 55, Appl	495	7	2.1	1959	3	US-08-993-722A-69	Sequence 69, Appl
C 423	7	2.1	1956	3	US-08-993-775B-57	Sequence 57, Appl	496	7	2.1	1959	3	US-08-993-722A-71	Sequence 71, App
C 424	7	2.1	1956	4	US-09-427-770-51	Sequence 51, Appl	497	7	2.1	1959	3	US-08-993-722A-73	Sequence 73, Appl1
C 425	7	2.1	1956	4	US-09-427-770-55	Sequence 55, Appl	498	7	2.1	1959	3	US-08-993-170A-5	Sequence 5, Appl1
C 426	7	2.1	1956	4	US-09-427-770-57	Sequence 57, Appl	499	7	2.1	1959	3	US-08-993-170A-7	Sequence 7, Appl1
C 427	7	2.1	1956	4	US-09-427-769-51	Sequence 51, Appl	500	7	2.1	1959	3	US-08-993-170A-9	Sequence 9, Appl1
C 428	7	2.1	1956	4	US-09-427-769-55	Sequence 55, Appl	501	7	2.1	1959	3	US-08-993-170A-11	Sequence 11, Appl
C 429	7	2.1	1956	4	US-09-427-769-57	Sequence 57, Appl	502	7	2.1	1959	3	US-08-993-170A-13	Sequence 13, Appl
C 430	7	2.1	1959	3	US-08-996-441B-1	Sequence 1, Appl1	503	7	2.1	1959	3	US-08-993-170A-15	Sequence 15, Appl
C 431	7	2.1	1959	3	US-08-996-441B-5	Sequence 5, Appl1	504	7	2.1	1959	3	US-08-993-170A-17	Sequence 17, Appl
C 432	7	2.1	1959	3	US-08-996-441B-7	Sequence 7, Appl1	505	7	2.1	1959	3	US-08-993-170A-19	Sequence 19, Appl
C 433	7	2.1	1959	3	US-08-996-441B-9	Sequence 9, Appl1	506	7	2.1	1959	3	US-08-993-170A-21	Sequence 21, Appl
C 434	7	2.1	1959	3	US-08-996-441B-11	Sequence 11, Appl	507	7	2.1	1959	3	US-08-993-170A-23	Sequence 23, Appl
C 435	7	2.1	1959	3	US-08-996-441B-13	Sequence 13, Appl	508	7	2.1	1959	3	US-08-993-170A-25	Sequence 25, Appl
C 436	7	2.1	1959	3	US-08-996-441B-15	Sequence 15, Appl	509	7	2.1	1959	3	US-08-993-170A-27	Sequence 27, Appl
C 437	7	2.1	1959	3	US-08-996-441B-17	Sequence 17, Appl	510	7	2.1	1959	3	US-08-993-170A-29	Sequence 29, Appl
C 438	7	2.1	1959	3	US-08-996-441B-19	Sequence 19, Appl	511	7	2.1	1959	3	US-08-993-170A-31	Sequence 31, Appl
C 439	7	2.1	1959	3	US-08-996-441B-21	Sequence 21, Appl	512	7	2.1	1959	3	US-08-993-170A-33	Sequence 33, Appl
C 440	7	2.1	1959	3	US-08-996-441B-23	Sequence 23, Appl	513	7	2.1	1959	3	US-08-993-170A-35	Sequence 35, Appl
C 441	7	2.1	1959	3	US-08-996-441B-25	Sequence 25, Appl	514	7	2.1	1959	3	US-08-993-170A-37	Sequence 37, Appl
C 442	7	2.1	1959	3	US-08-996-441B-27	Sequence 27, Appl	515	7	2.1	1959	3	US-08-993-170A-39	Sequence 39, Appl
C 443	7	2.1	1959	3	US-08-996-441B-29	Sequence 29, Appl	516	7	2.1	1959	3	US-08-993-170A-41	Sequence 41, Appl
C 444	7	2.1	1959	3	US-08-996-441B-31	Sequence 31, Appl	517	7	2.1	1959	3	US-08-993-170A-43	Sequence 43, Appl
C 445	7	2.1	1959	3	US-08-996-441B-33	Sequence 33, Appl	518	7	2.1	1959	3	US-08-993-170A-45	Sequence 45, Appl
C 446	7	2.1	1959	3	US-08-996-441B-35	Sequence 35, Appl	519	7	2.1	1959	3	US-08-993-170A-47	Sequence 47, Appl
C 447	7	2.1	1959	3	US-08-996-441B-37	Sequence 37, Appl	520	7	2.1	1959	3	US-08-993-170A-49	Sequence 49, Appl
C 448	7	2.1	1959	3	US-08-996-441B-39	Sequence 39, Appl	521	7	2.1	1959	3	US-08-993-170A-51	Sequence 51, Appl
C 449	7	2.1	1959	3	US-08-996-441B-41	Sequence 41, Appl	522	7	2.1	1959	3	US-08-993-170A-53	Sequence 53, Appl
C 450	7	2.1	1959	3	US-08-996-441B-43	Sequence 43, Appl	523	7	2.1	1959	3	US-08-993-170A-55	Sequence 55, Appl

524	7	2.1	1959	3	US-08-993-170A-63	Sequence 63, Appl	597	7	2.1	1959	4	US-09-427-769-1	Sequence 1, Appl
525	7	2.1	1959	3	US-08-993-170A-65	Sequence 65, Appl	598	7	2.1	1959	4	US-09-427-769-3	Sequence 3, Appl
526	7	2.1	1959	3	US-08-993-170A-67	Sequence 67, Appl	599	7	2.1	1959	4	US-09-427-769-5	Sequence 5, Appl
527	7	2.1	1959	3	US-08-993-170A-69	Sequence 69, Appl	600	7	2.1	1959	4	US-09-427-769-7	Sequence 7, Appl
528	7	2.1	1959	3	US-08-993-170A-97	Sequence 97, Appl	601	7	2.1	1959	4	US-09-427-769-9	Sequence 9, Appl
529	7	2.1	1959	3	US-08-993-170A-107	Sequence 107, App	602	7	2.1	1959	4	US-09-427-769-11	Sequence 11, Appl
530	7	2.1	1959	3	US-08-993-775B-1	Sequence 1, Appl	603	7	2.1	1959	4	US-09-427-769-13	Sequence 13, Appl
531	7	2.1	1959	3	US-08-993-775B-3	Sequence 3, Appl	604	7	2.1	1959	4	US-09-427-769-15	Sequence 15, Appl
532	7	2.1	1959	3	US-08-993-775B-5	Sequence 5, Appl	605	7	2.1	1959	4	US-09-427-769-17	Sequence 17, Appl
533	7	2.1	1959	3	US-08-993-775B-7	Sequence 7, Appl	606	7	2.1	1959	4	US-09-427-769-19	Sequence 19, Appl
534	7	2.1	1959	3	US-08-993-775B-9	Sequence 9, Appl	607	7	2.1	1959	4	US-09-427-769-21	Sequence 21, Appl
535	7	2.1	1959	3	US-08-993-775B-11	Sequence 11, Appl	608	7	2.1	1959	4	US-09-427-769-23	Sequence 23, Appl
536	7	2.1	1959	3	US-08-993-775B-13	Sequence 13, Appl	609	7	2.1	1959	4	US-09-427-769-25	Sequence 25, Appl
537	7	2.1	1959	3	US-08-993-775B-15	Sequence 15, Appl	610	7	2.1	1959	4	US-09-427-769-27	Sequence 27, Appl
538	7	2.1	1959	3	US-08-993-775B-17	Sequence 17, Appl	611	7	2.1	1959	4	US-09-427-769-29	Sequence 29, Appl
539	7	2.1	1959	3	US-08-993-775B-19	Sequence 19, Appl	612	7	2.1	1959	4	US-09-427-769-31	Sequence 31, Appl
540	7	2.1	1959	3	US-08-993-775B-21	Sequence 21, Appl	613	7	2.1	1959	4	US-09-427-769-33	Sequence 33, Appl
541	7	2.1	1959	3	US-08-993-775B-23	Sequence 23, Appl	614	7	2.1	1959	4	US-09-427-769-35	Sequence 35, Appl
542	7	2.1	1959	3	US-08-993-775B-25	Sequence 25, Appl	615	7	2.1	1959	4	US-09-427-769-37	Sequence 37, Appl
543	7	2.1	1959	3	US-08-993-775B-27	Sequence 27, Appl	616	7	2.1	1959	4	US-09-427-769-39	Sequence 39, Appl
544	7	2.1	1959	3	US-08-993-775B-29	Sequence 29, Appl	617	7	2.1	1959	4	US-09-427-769-41	Sequence 41, Appl
545	7	2.1	1959	3	US-08-993-775B-31	Sequence 31, Appl	618	7	2.1	1959	4	US-09-427-769-43	Sequence 43, Appl
546	7	2.1	1959	3	US-08-993-775B-33	Sequence 33, Appl	619	7	2.1	1959	4	US-09-427-769-45	Sequence 45, Appl
547	7	2.1	1959	3	US-08-993-775B-35	Sequence 35, Appl	620	7	2.1	1959	4	US-09-427-769-47	Sequence 47, Appl
548	7	2.1	1959	3	US-08-993-775B-37	Sequence 37, Appl	621	7	2.1	1959	4	US-09-427-769-49	Sequence 49, Appl
549	7	2.1	1959	3	US-08-993-775B-39	Sequence 39, Appl	622	7	2.1	1959	4	US-09-427-769-51	Sequence 51, Appl
550	7	2.1	1959	3	US-08-993-775B-41	Sequence 41, Appl	623	7	2.1	1959	4	US-09-427-769-53	Sequence 53, Appl
551	7	2.1	1959	3	US-08-993-775B-43	Sequence 43, Appl	624	7	2.1	1959	4	US-09-427-769-55	Sequence 55, Appl
552	7	2.1	1959	3	US-08-993-775B-45	Sequence 45, Appl	625	7	2.1	1959	4	US-09-427-769-57	Sequence 57, Appl
553	7	2.1	1959	3	US-08-993-775B-47	Sequence 47, Appl	626	7	2.1	1959	4	US-09-427-769-59	Sequence 59, Appl
554	7	2.1	1959	3	US-08-993-775B-49	Sequence 49, Appl	627	7	2.1	1959	4	US-09-427-769-61	Sequence 61, Appl
555	7	2.1	1959	3	US-08-993-775B-51	Sequence 51, Appl	628	7	2.1	1959	4	US-09-427-769-63	Sequence 63, Appl
556	7	2.1	1959	3	US-08-993-775B-53	Sequence 53, Appl	629	7	2.1	1959	4	US-09-427-769-65	Sequence 65, Appl
557	7	2.1	1959	3	US-08-993-775B-55	Sequence 55, Appl	630	7	2.1	1959	4	US-09-427-769-67	Sequence 67, Appl
558	7	2.1	1959	3	US-08-993-775B-57	Sequence 57, Appl	631	7	2.1	1962	4	US-09-427-769-69	Sequence 69, Appl
559	7	2.1	1959	3	US-08-993-775B-59	Sequence 59, Appl	C 631	7	2.1	1971	4	US-09-377-466B-5	Sequence 107, App
560	7	2.1	1959	3	US-08-993-775B-61	Sequence 61, Appl	632	7	2.1	1972	6	US-09-107-532A-3639	Sequence 3639, Ap
561	7	2.1	1959	3	US-08-993-775B-63	Sequence 63, Appl	633	7	2.1	1972	6	US-09-107-532A-3639	Sequence 3639, Ap
562	7	2.1	1959	3	US-08-993-775B-65	Sequence 65, Appl	634	7	2.1	1972	6	5187091-1	Patent No. 5187091
563	7	2.1	1959	4	US-09-377-466B-1	Sequence 1, Appl	635	7	2.1	1984	4	US-09-377-466B-9	Sequence 9, Appl
564	7	2.1	1959	4	US-09-427-770-1	Sequence 1, Appl	636	7	2.1	1984	4	US-09-377-466B-11	Sequence 11, Appl
565	7	2.1	1959	4	US-09-427-770-3	Sequence 3, Appl	637	7	2.1	1989	4	US-08-845-258-4	Sequence 7, Appl
566	7	2.1	1959	4	US-09-427-770-5	Sequence 5, Appl	C 637	7	2.1	1991	3	US-08-845-258-4	Sequence 4, Appl
567	7	2.1	1959	4	US-09-427-770-7	Sequence 7, Appl	C 638	7	2.1	1991	3	US-08-990-571-4	Sequence 4, Appl
568	7	2.1	1959	4	US-09-427-770-9	Sequence 9, Appl	C 639	7	2.1	1991	3	US-08-723-142A-4	Sequence 4, Appl
569	7	2.1	1959	4	US-09-427-770-11	Sequence 11, Appl	C 640	7	2.1	1991	3	US-09-528-784A-4	Sequence 4, Appl
570	7	2.1	1959	4	US-09-427-770-13	Sequence 13, Appl	C 641	7	2.1	1996	4	US-09-569-098A-4	Sequence 4, Appl
571	7	2.1	1959	4	US-09-427-770-15	Sequence 15, Appl	642	7	2.1	1996	4	US-09-949-016-5085	Sequence 5085, Ap
572	7	2.1	1959	4	US-09-427-770-17	Sequence 17, Appl	643	7	2.1	2000	3	US-08-996-441B-99	Sequence 99, Appl
573	7	2.1	1959	4	US-09-427-770-19	Sequence 19, Appl	644	7	2.1	2000	3	US-08-993-722A-99	Sequence 99, Appl
574	7	2.1	1959	4	US-09-427-770-21	Sequence 21, Appl	645	7	2.1	2000	3	US-08-993-775B-99	Sequence 99, Appl
575	7	2.1	1959	4	US-09-427-770-23	Sequence 23, Appl	646	7	2.1	2000	3	US-09-427-770-99	Sequence 99, Appl
576	7	2.1	1959	4	US-09-427-770-25	Sequence 25, Appl	647	7	2.1	2000	4	US-09-427-769-99	Sequence 99, Appl
577	7	2.1	1959	4	US-09-427-770-27	Sequence 27, Appl	648	7	2.1	2004	4	US-09-710-279-197	Sequence 197, App
578	7	2.1	1959	4	US-09-427-770-29	Sequence 29, Appl	C 649	7	2.1	2004	4	US-09-643-597-153	Sequence 153, App
579	7	2.1	1959	4	US-09-427-770-31	Sequence 31, Appl	650	7	2.1	2007	3	US-09-480-884A-153	Sequence 153, App
580	7	2.1	1959	4	US-09-427-770-33	Sequence 33, Appl	651	7	2.1	2007	4	US-09-542-615A-153	Sequence 153, App
581	7	2.1	1959	4	US-09-427-770-35	Sequence 35, Appl	652	7	2.1	2007	4	US-09-606-421B-153	Sequence 153, App
582	7	2.1	1959	4	US-09-427-770-37	Sequence 37, Appl	653	7	2.1	2007	4	US-09-221-107-153	Sequence 153, App
583	7	2.1	1959	4	US-09-427-770-39	Sequence 39, Appl	654	7	2.1	2007	4	US-09-466-396A-153	Sequence 153, App
584	7	2.1	1959	4	US-09-427-770-41	Sequence 41, Appl	655	7	2.1	2007	4	US-09-476-496A-153	Sequence 153, App
585	7	2.1	1959	4	US-09-427-770-43	Sequence 43, Appl	656	7	2.1	2007	4	US-09-630-940B-153	Sequence 153, App
586	7	2.1	1959	4	US-09-427-770-45	Sequence 45, Appl	657	7	2.1	2007	4	US-09-285-479-153	Sequence 153, App
587	7	2.1	1959	4	US-09-427-770-47	Sequence 47, Appl	658	7	2.1	2041	4	US-09-566-921-110	Sequence 110, App
588	7	2.1	1959	4	US-09-427-770-49	Sequence 49, Appl	659	7	2.1	2042	4	US-09-673-395A-91	Sequence 91, Appl
589	7	2.1	1959	4	US-09-427-770-51	Sequence 51, Appl	660	7	2.1	2042	4	US-09-673-395A-105	Sequence 105, App
590	7	2.1	1959	4	US-09-427-770-53	Sequence 53, Appl	661	7	2.1	2043	4	US-09-583-110-530	Sequence 530, App
591	7	2.1	1959	4	US-09-427-770-55	Sequence 55, Appl	662	7	2.1	2050	3	US-08-996-441B-101	Sequence 101, App
592	7	2.1	1959	4	US-09-427-770-57	Sequence 57, Appl	663	7	2.1	2050	3	US-08-993-722A-101	Sequence 101, App
593	7	2.1	1959	4	US-09-427-770-59	Sequence 59, Appl	664	7	2.1	2050	3	US-08-993-170A-101	Sequence 101, App
594	7	2.1	1959	4	US-09-427-770-61	Sequence 61, Appl	665	7	2.1	2050	3	US-08-993-775B-101	Sequence 101, App
595	7	2.1	1959	4	US-09-427-770-63	Sequence 63, Appl	666	7	2.1	2050	4	US-09-427-770-101	Sequence 101, App
596	7	2.1	1959	4	US-09-427-770-65	Sequence 65, Appl	667	7	2.1	2050	4	US-09-427-769-101	Sequence 101, App
597	7	2.1	1959	4	US-09-427-770-67	Sequence 67, Appl	668	7	2.1	2053	4	US-09-107-433-232	Sequence 232, App
598	7	2.1	1959	4	US-09-427-770-69	Sequence 69, Appl	C 669	7	2.1	2073	4		
599	7	2.1	1959	4	US-09-427-770-71	Sequence 71, Appl							
600	7	2.1	1959	4	US-09-427-770-73	Sequence 73, Appl							
601	7	2.1	1959	4	US-09-427-770-75	Sequence 75, Appl							
602	7	2.1	1959	4	US-09-427-770-77	Sequence 77, Appl							
603	7	2.1	1959	4	US-09-427-770-79	Sequence 79, Appl							
604	7	2.1	1959	4	US-09-427-770-81	Sequence 81, Appl							
605	7	2.1	1959	4	US-09-427-770-83	Sequence 83, Appl							
606	7	2.1	1959	4	US-09-427-770-85	Sequence 85, Appl							
607	7	2.1	1959	4	US-09-427-770-87	Sequence 87, Appl							
608	7	2.1	1959	4	US-09-427-770-89	Sequence 89, Appl							
609	7	2.1	1959	4	US-09-427-770-91	Sequence 91, Appl							
610	7	2.1	1959	4	US-09-427-770-93	Sequence 93, Appl							
611	7	2.1	1959	4	US-09-427-770-95	Sequence 95, Appl							
612	7	2.1	1959	4	US-09-427-770-97	Sequence 97, Appl							
613	7	2.1	1959	4	US-09-427-770-99	Sequence 99, Appl							
614	7	2.1	1959	4	US-09-427-770-101	Sequence 101, Appl							
615	7	2.1	1959	4	US-09-427-770-103								

c 670	7	2.1	2091	4	US-09-620-312D-743	Sequence 743, App	c 743	7	2.1	3423	4	US-09-614-221A-91	Sequence 91, Appl
c 671	7	2.1	2097	3	US-09-134-001C-795	Sequence 795, App	c 744	7	2.1	3438	4	US-10-164-595-29	Sequence 29, Appl
c 672	7	2.1	2121	3	US-09-134-001C-2071	Sequence 2071, Ap	c 745	7	2.1	3450	4	US-09-377-466B-17	Sequence 17, Appl
c 673	7	2.1	2122	4	US-09-710-279-4457	Sequence 4457, Ap	c 746	7	2.1	3455	4	US-09-377-466B-36	Sequence 36, Appl
c 674	7	2.1	2126	4	US-09-620-312D-598	Sequence 598, App	c 747	7	2.1	3459	4	US-09-710-279-3553	Sequence 3553, Ap
c 675	7	2.1	2148	3	US-09-643-597-154	Sequence 154, App	c 748	7	2.1	3469	4	US-09-377-466B-23	Sequence 23, Appl
c 676	7	2.1	2148	4	US-09-480-884A-154	Sequence 154, App	c 749	7	2.1	3513	4	US-09-815-923-13	Sequence 13, Appl
c 677	7	2.1	2148	4	US-09-542-615A-154	Sequence 154, App	c 750	7	2.1	3522	4	US-09-543-681A-3793	Sequence 3793, Ap
c 678	7	2.1	2148	4	US-09-606-421B-154	Sequence 154, App	c 751	7	2.1	3532	2	US-09-205-204-1	Sequence 1, Appl1
c 679	7	2.1	2148	4	US-09-221-107-154	Sequence 154, App	c 752	7	2.1	3549	4	US-09-710-279-3796	Sequence 3796, Ap
c 680	7	2.1	2148	4	US-09-476-496A-154	Sequence 154, App	c 753	7	2.1	3618	4	US-09-710-279-3443	Sequence 3443, Ap
c 681	7	2.1	2148	4	US-09-630-940B-154	Sequence 154, App	c 754	7	2.1	3627	4	US-09-710-279-3764	Sequence 3764, Ap
c 682	7	2.1	2148	4	US-09-385-479-154	Sequence 154, App	c 755	7	2.1	3704	4	US-10-160-719A-57	Sequence 57, Appl
c 683	7	2.1	2148	4	US-09-086-663A-78	Sequence 78, Appl	c 756	7	2.1	3732	3	US-09-212-971-7	Sequence 7, Appl1
c 684	7	2.1	2156	4	US-08-973-005A-10	Sequence 10, Appl	c 757	7	2.1	3732	3	US-08-800-929A-7	Sequence 7, Appl1
c 685	7	2.1	2157	3	US-09-949-016-3018	Sequence 3018, Ap	c 758	7	2.1	3732	3	US-09-617-053A-7	Sequence 7, Appl1
c 686	7	2.1	2158	4	US-09-949-016-3026	Sequence 3026, Ap	c 759	7	2.1	3754	4	US-09-377-466B-15	Sequence 15, Appl
c 687	7	2.1	2158	4	US-09-949-016-3026	Sequence 3026, Ap	c 760	7	2.1	3883	1	US-08-468-036-33	Sequence 33, Appl
c 688	7	2.1	2172	3	US-09-134-001C-2223	Sequence 2223, Ap	c 761	7	2.1	3883	2	US-08-376-843-33	Sequence 33, Appl
c 689	7	2.1	2202	4	US-09-902-540-8501	Sequence 8501, Ap	c 762	7	2.1	3884	3	US-09-541-782-3	Sequence 3, Appl1
c 690	7	2.1	2253	4	US-09-949-016-3908	Sequence 3908, Ap	c 763	7	2.1	3884	3	US-10-270-085-3	Sequence 6, Appl1
c 691	7	2.1	2256	4	US-09-949-016-2427	Sequence 2427, Ap	c 764	7	2.1	3901	2	US-10-270-085-3	Sequence 6, Appl1
c 692	7	2.1	2256	4	US-09-949-016-2428	Sequence 2428, Ap	c 765	7	2.1	3901	3	US-08-574-959A-6	Sequence 4309, Ap
c 693	7	2.1	2280	3	US-08-996-441B-102	Sequence 102, App	c 766	7	2.1	3901	3	US-09-357-014-6	Sequence 816, App
c 694	7	2.1	2280	3	US-08-993-722A-102	Sequence 102, App	c 767	7	2.1	3983	4	US-09-710-279-4309	Sequence 816, App
c 695	7	2.1	2280	3	US-08-993-775B-102	Sequence 102, App	c 768	7	2.1	3989	4	US-08-956-171B-816	Sequence 816, App
c 696	7	2.1	2280	4	US-09-427-770-102	Sequence 102, App	c 769	7	2.1	3989	4	US-08-781-986A-816	Sequence 969, App
c 697	7	2.1	2280	4	US-09-427-769-102	Sequence 102, App	c 770	7	2.1	4121	4	US-09-023-655-969	Sequence 229, App
c 698	7	2.1	2280	4	US-09-540-236-862	Sequence 862, App	c 771	7	2.1	4146	4	US-09-614-221A-229	Sequence 13, Appl
c 699	7	2.1	2310	4	US-09-540-236-862	Sequence 862, App	c 772	7	2.1	4196	4	US-09-377-466B-13	Sequence 236, App
c 700	7	2.1	2334	4	US-09-248-796A-1911	Sequence 1911, Ap	c 773	7	2.1	4209	4	US-09-774-528-236	Sequence 400, App
c 701	7	2.1	2364	2	US-08-981-690-1	Sequence 1, Appli	c 774	7	2.1	4215	4	US-09-248-796A-400	Sequence 4000, Ap
c 702	7	2.1	2370	4	US-09-252-991A-9040	Sequence 9040, Ap	c 775	7	2.1	4215	4	US-09-669-980A-4000	Sequence 6, Appl1
c 703	7	2.1	2379	4	US-09-252-991A-9252	Sequence 9252, Ap	c 776	7	2.1	4310	4	US-09-773-426A-4	Sequence 4, Appl1
c 704	7	2.1	2379	4	US-09-543-681A-2445	Sequence 2445, Ap	c 777	7	2.1	4321	4	US-10-314-881-4	Sequence 4, Appl1
c 705	7	2.1	2406	2	US-08-687-702-2	Sequence 2, Appli	c 778	7	2.1	4321	4	US-09-495-823-4	Sequence 546, App
c 706	7	2.1	2430	5	PCT-US92-00040-1	Sequence 4, Appli	c 780	7	2.1	4336	4	US-08-956-171E-546	Sequence 546, App
c 707	7	2.1	2450	4	US-09-618-425-4	Sequence 223, App	c 781	7	2.1	4336	4	US-08-781-986A-546	Sequence 6, Appl1
c 708	7	2.1	2494	4	US-08-239-276-7	Sequence 7, Appli	c 782	7	2.1	4528	4	US-08-809-254A-6	Sequence 15228, A
c 709	7	2.1	2599	2	US-08-468-579B-7	Sequence 7, Appli	c 783	7	2.1	4587	4	US-09-949-016-15228	Sequence 9, Appl1
c 710	7	2.1	2599	3	US-08-468-577B-7	Sequence 7, Appli	c 784	7	2.1	4621	4	US-09-949-016-3235	Sequence 3235, Ap
c 711	7	2.1	2599	3	US-08-468-581A-7	Sequence 7, Appli	c 785	7	2.1	4767	4	US-09-231-899-76	Sequence 76, Appl
c 712	7	2.1	2599	3	US-08-468-581A-7	Sequence 7, Appli	c 786	7	2.1	4767	4	US-09-561-497-10	Sequence 10, Appl
c 713	7	2.1	2632	4	US-09-205-258-110	Sequence 110, App	c 787	7	2.1	4793	3	US-09-270-767-10504	Sequence 10504, A
c 714	7	2.1	2694	4	US-09-902-540-4257	Sequence 4257, Ap	c 788	7	2.1	4860	4	US-09-445-353E-1	Sequence 5, Appl1
c 715	7	2.1	2735	4	US-09-620-312D-842	Sequence 842, App	c 789	7	2.1	4899	2	US-08-210-535-5	Patent No. 5268270
c 716	7	2.1	2790	3	US-09-370-838-79	Sequence 79, Appl	c 790	7	2.1	4899	6	5268270-1	Patent No. 5268270
c 717	7	2.1	2790	3	US-09-854-133-79	Sequence 79, Appl	c 791	7	2.1	4899	6	5268270-1	Sequence 16010, A
c 718	7	2.1	2995	1	US-08-592-136-85	Sequence 85, Appl	c 792	7	2.1	5164	4	US-09-949-016-16010	Sequence 613, App
c 719	7	2.1	2995	1	US-08-592-136-85	Sequence 85, Appl	c 793	7	2.1	5203	4	US-09-799-451-613	Sequence 13950, A
c 720	7	2.1	3036	4	US-09-168-595-85	Sequence 3822, Ap	c 794	7	2.1	5203	4	US-09-949-016-13950	Sequence 71, Appl
c 721	7	2.1	3039	4	US-09-710-279-3822	Sequence 19, Appl	c 795	7	2.1	5215	4	US-09-231-899-71	Sequence 4074, Ap
c 722	7	2.1	3039	4	US-09-377-466B-19	Sequence 21, Appl	c 796	7	2.1	5226	4	US-09-583-110-1314	Sequence 1314, Ap
c 723	7	2.1	3044	4	US-09-377-466B-38	Sequence 38, Appl	c 797	7	2.1	5304	4	US-09-583-110-1314	Sequence 1073, Ap
c 724	7	2.1	3063	4	US-09-620-312D-474	Sequence 474, App	c 798	7	2.1	5310	4	US-09-107-433-1073	Sequence 1588, Ap
c 725	7	2.1	3074	4	US-09-270-767-13401	Sequence 13401, A	c 799	7	2.1	5337	1	US-09-252-991A-1588	Sequence 15, Appl
c 726	7	2.1	3084	1	US-08-551-437-3	Sequence 3, Appli	c 800	7	2.1	5348	4	US-08-441-139-15	Sequence 1259, Ap
c 727	7	2.1	3084	3	US-09-004-225-3	Sequence 3, Appli	c 801	7	2.1	5522	4	US-09-949-016-1259	Sequence 1014, Ap
c 728	7	2.1	3084	3	US-09-084-346-3	Sequence 3, Appli	c 802	7	2.1	5523	4	US-09-949-016-1014	Sequence 1, Appli
c 729	7	2.1	3084	4	US-09-902-540-536	Sequence 536, App	c 803	7	2.1	5720	3	US-09-442-100-1	Sequence 1, Appli
c 730	7	2.1	3084	4	US-08-943-768-1	Sequence 1, Appli	c 804	7	2.1	5720	4	US-08-939-106-1	Sequence 1, Appl1
c 731	7	2.1	3150	3	US-08-865-960-1	Sequence 1, Appli	c 805	7	2.1	5720	4	US-09-442-102-1	Sequence 183, App
c 732	7	2.1	3150	3	US-08-050-684-1	Sequence 1, Appli	c 806	7	2.1	5933	4	US-09-976-594-183	Sequence 17300, A
c 733	7	2.1	3175	1	US-08-582-719-1	Sequence 1, Appli	c 807	7	2.1	5939	4	US-09-949-016-17300	Sequence 12512, A
c 734	7	2.1	3175	1	US-08-582-719-1	Sequence 1, Appli	c 808	7	2.1	5941	4	US-09-949-016-12512	Sequence 869, App
c 735	7	2.1	3181	2	US-09-710-279-4118	Sequence 4118, Ap	c 809	7	2.1	6019	4	US-09-949-016-869	Sequence 2671, Ap
c 736	7	2.1	3211	3	US-08-574-959A-8	Sequence 8, Appli	c 810	7	2.1	6019	4	US-09-949-016-2671	Sequence 3207, Ap
c 737	7	2.1	3211	3	US-09-357-014-8	Sequence 8, Appli	c 811	7	2.1	6165	3	US-09-134-000C-3207	Sequence 457, App
c 738	7	2.1	3236	4	US-09-710-279-3492	Sequence 3492, Ap	c 812	7	2.1	6168	3	US-09-071-035-457	Sequence 461, App
c 739	7	2.1	3297	4	US-09-620-312D-417	Sequence 417, App	c 813	7	2.1	6168	3	US-09-071-035-461	Sequence 465, App
c 740	7	2.1	3376	3	US-08-320-559-29	Sequence 29, Appl	c 814	7	2.1	6168	3	US-09-071-035-461	Sequence 465, App
c 741	7	2.1	3376	3	US-08-545-860D-29	Sequence 29, Appl	c 815	7	2.1	6168	3	US-09-071-035-465	Sequence 465, App
c 742	7	2.1	3376	5	PCT-US94-04496-29	Sequence 29, Appl							





US-10-724-972A-6352 (1-335) x US-09-134-001C-2023 (1-183)

QY 285 ProValTriPlyLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgApp 304  
Db 28 CTGTATGGAAGAAATTAACCGCTGTGAAAAATCAACGTTGTGATATTTAGACCGTGAC 87  
QY 305 LeuTriPalaArSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGlu 324  
Db 88 TTATGGCAAGATCACGTGGTTTAATTTCTCAGAGAAATGCAAAAGAACTTGTGTGAA 147

QY 325 LeuSerLysLysAspSerLysLysAspAsnLys 335  
Db 148 TTATCTAAGAAGATAGTAAAAAGATAATAAG 180

RESULT 2  
US-09-710-279-3782  
; Sequence 3782, Application US/09710279  
; Patent No. 6703492  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/09/710,279  
; CURRENT FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3782  
; LENGTH: 3046  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-09-710-279-3782

Alignment Scores:  
Pred. No.: 6.46e-26 Length: 3046  
Score: 35.00 Matches: 35  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.45% Indels: 0  
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-710-279-3782 (1-3046)

QY 1 GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuPheValLeu 20  
Db 2705 GGAGTGAATCAGTCAGAGGTTTAAAAATTTTAAAGTGAATGGCTTATTGTTGTTTA 2764

QY 21 IleAlaThrAlaAlaCysGlyAsnSerSerSerAsnSerSer 35  
Db 2765 ATTGCACTGCAGCATGTGGAATATATATCTCAAGTAACTCAAGT 2809

RESULT 3  
US-08-956-171E-2556  
; Sequence 2556, Application US/08956171E  
; Patent No. 6593114  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; Gil H. Choi  
; Patrick S. Dillon  
; Craig A. Rosen  
; Steven C. Barash  
; Michael R. Fannon  
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5256  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland

Sequence 14706, A  
Sequence 14707, A  
Sequence 14708, A  
Sequence 14709, A  
Sequence 15854, A  
Sequence 12007, A  
Sequence 16781, A  
Sequence 12838, A  
Sequence 17458, A  
Sequence 11796, A  
Sequence 12843, A  
Sequence 12844, A  
Sequence 12846, A  
Sequence 13542, A  
Sequence 13543, A  
Sequence 13544, A  
Sequence 13545, A  
Sequence 14633, A  
Sequence 14634, A  
Sequence 14635, A  
Sequence 14636, A  
Sequence 14637, A  
Sequence 14638, A  
Sequence 14639, A  
Sequence 14640, A  
Sequence 12232, A  
Sequence 30, Appl  
Sequence 12150, A  
Sequence 15374, A  
Sequence 16910, A  
Sequence 12175, A  
Sequence 12638, A  
Sequence 16057, A  
Sequence 16076, A  
Sequence 14712, A  
Sequence 14713, A  
Sequence 11803, A  
Sequence 16739, A  
Sequence 14817, A

US-09-949-016-14706  
US-09-949-016-14707  
US-09-949-016-14708  
US-09-949-016-14709  
US-09-949-016-15854  
US-09-949-016-12007  
US-09-949-016-16781  
US-09-949-016-12838  
US-09-949-016-17458  
US-09-949-016-11796  
US-09-949-016-12843  
US-09-949-016-12844  
US-09-949-016-12846  
US-09-949-016-13542  
US-09-949-016-13543  
US-09-949-016-13544  
US-09-949-016-13545  
US-09-949-016-14633  
US-09-949-016-14634  
US-09-949-016-14635  
US-09-949-016-14636  
US-09-949-016-14637  
US-09-949-016-14638  
US-09-949-016-14639  
US-09-949-016-14640  
US-09-949-016-12232  
US-09-596-002-30  
US-09-949-016-12150  
US-09-949-016-15374  
US-09-949-016-16910  
US-09-949-016-12175  
US-09-949-016-12638  
US-09-949-016-16057  
US-09-949-016-16076  
US-09-949-016-14712  
US-09-949-016-14713  
US-09-949-016-11803  
US-09-949-016-16739  
US-09-949-016-14817

ALIGNMENTS

RESULT 1  
US-09-134-001C-2023  
; Sequence 2023, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 2023  
; LENGTH: 183  
; TYPE: DNA  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-2023

Alignment Scores:  
Pred. No.: 1.61e-43 Length: 183  
Score: 51.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 15.22% Indels: 0  
DB: 3 Gaps: 0

; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; FILING DATE: 20-Oct-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/009,861  
; FILING DATE: January 5, 1996  
; APPLICATION NUMBER: 08/781,986  
; FILING DATE: January 3, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mark J. Hyman  
; REGISTRATION NUMBER: 46,789  
; REFERENCE/DOCKET NUMBER: PB248P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (240) 314-1224  
; TELEFAX: (301) 309-8439  
; INFORMATION FOR SEQ ID NO: 2556:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 242 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; US-08-956-171E-2556

Alignment Scores:  
Pred. No.: 1,01e-17 Length: 242  
Score: 26.00 Matches: 26  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 7.76% Indels: 0  
DB: 4 Gaps: 0  
US-10-724-972A-6352 (1-335) x US-08-956-171E-2556 (1-242)

QY 102 TyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluSerLysLeuLysPro 121  
Db 98 TATACCTTCTGTAGGTACACGTAACAGCCAACTTAGAAGAAATTAGTAAATTAACCG 157  
QY 122 AspLeuIleIleAlaAsp 127  
Db 158 GATTTAATTATCGCTGAT 175

RESULT 4  
US-08-781-986A-2556  
; Sequence 2556, Application US/08781986A  
; Patent No. 6737248  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5255  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/781,986A  
; FILING DATE:

; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Benson, Bob  
; REGISTRATION NUMBER: 30,446  
; REFERENCE/DOCKET NUMBER: PB248PP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 2556:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 242 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; US-08-781-986A-2556

Alignment Scores:  
Pred. No.: 1,01e-17 Length: 242  
Score: 26.00 Matches: 26  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 7.76% Indels: 0  
DB: 4 Gaps: 0  
US-10-724-972A-6352 (1-335) x US-08-781-986A-2556 (1-242)

QY 102 TyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluSerLysLeuLysPro 121  
Db 98 TATACCTTCTGTAGGTACACGTAACAGCCAACTTAGAAGAAATTAGTAAATTAACCG 157  
QY 122 AspLeuIleIleAlaAsp 127  
Db 158 GATTTAATTATCGCTGAT 175

## RESULT 5

US-08-956-171E-355  
; Sequence 355, Application US/08956171E  
; Patent No. 6593114  
; GENERAL INFORMATION:

APPLICANT: Charles Kunsch  
Gil H. Choi  
Patrick S. Dillon  
Craig A. Rosen  
Steven C. Barash  
Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 5256  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/956,171E  
FILING DATE: 20-Oct-1997

CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,861  
FILING DATE: January 5, 1996  
APPLICATION NUMBER: 08/781,986  
FILING DATE: January 3, 1997  
ATTORNEY/AGENT INFORMATION:

NAME: Mark J. Hyman

REGISTRATION NUMBER: 46,789  
REFERENCE/DOCKET NUMBER: PB248P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (240) 314-1224  
TELEFAX: (301) 309-8439  
INFORMATION FOR SEQ ID NO: 355:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 668 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 355:  
US-08-956-171E-355

Alignment Scores:  
Pred. No.: 3,21e-14 Length: 668  
Score: 23.00 Matches: 23  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.87% Indels: 0  
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-08-781-986A-355 (1-668)

QY 306 TtpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeu 325  
Db 290 TGGGCAAGATCTCGTGGCTTAATTTCTCTGAAGAAATGCTTAAGAACTTGTGAATT 349

QY 326 SerLysLys 328  
Db 350 TCAAAAAA 358

RESULT 7  
US-09-830-217-15  
Sequence 15, Application US/09830217  
Patent No. 6521441  
GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides  
FILE REFERENCE: PB461PCT  
CURRENT APPLICATION NUMBER: US/09/830,217  
CURRENT FILING DATE: 2001-04-24  
PRIOR APPLICATION NUMBER: PCT/US99/06199  
PRIOR FILING DATE: 1999-03-18  
PRIOR APPLICATION NUMBER: 60/080,296  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/084,674  
PRIOR FILING DATE: 1998-05-07  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 15  
LENGTH: 999  
TYPE: DNA  
ORGANISM: Staphylococcus aureus  
US-09-830-217-15

Alignment Scores:  
Pred. No.: 8,28e-05 Length: 999  
Score: 14.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.18% Indels: 0  
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-830-217-15 (1-999)

QY 111 ProAenLeuGluGluIleSerLysLeuLysProAepLeuIle 124  
Db 322 CCTAACTTAGAGGAAATCTTAATAATTAACCGGACTTAATT 363

RESULT 8  
US-10-278-946-15  
Sequence 15, Application US/10278946  
Patent No. 6821754  
GENERAL INFORMATION:  
APPLICANT: Simpson et al.  
TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides  
FILE REFERENCE: PB461USDI  
CURRENT APPLICATION NUMBER: US/10/278,946  
CURRENT FILING DATE: 2002-10-24  
PRIOR APPLICATION NUMBER: US 09/830,217  
PRIOR FILING DATE: 2001-04-24  
PRIOR APPLICATION NUMBER: PCT/US99/06199  
PRIOR FILING DATE: 1999-03-18  
PRIOR APPLICATION NUMBER: US 60/078,682  
PRIOR FILING DATE: 1998-03-20

REGISTRATION NUMBER: 46,789  
REFERENCE/DOCKET NUMBER: PB248P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (240) 314-1224  
TELEFAX: (301) 309-8439  
INFORMATION FOR SEQ ID NO: 355:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 668 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 355:  
US-08-956-171E-355

Alignment Scores:  
Pred. No.: 3,21e-14 Length: 668  
Score: 23.00 Matches: 23  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.87% Indels: 0  
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-08-956-171E-355 (1-668)

QY 306 TtpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeu 325  
Db 290 TGGGCAAGATCTCGTGGCTTAATTTCTCTGAAGAAATGCTTAAGAACTTGTGAATT 349

QY 326 SerLysLys 328  
Db 350 TCAAAAAA 358

RESULT 6  
US-08-781-986A-355  
Sequence 355, Application US/08781986A  
Patent No. 6737248  
GENERAL INFORMATION:  
APPLICANT: Charles Kunsch  
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 5255  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/781,986A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Benson, Bob  
REGISTRATION NUMBER: 30,446  
REFERENCE/DOCKET NUMBER: PB248PP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 355:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 668 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-781-986A-355

SEQUENCE DESCRIPTION: SEQ ID NO: 238:  
US-08-956-171E-238  
Alignment Scores:  
Pred. No.: 0.000295 Length: 3775  
Score: 14.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.18% Indels: 0  
DB: 4 Gaps: 0  
US-10-724-972A-6352 (1-335) x US-08-956-171E-238 (1-3775)  
QY 111 ProAsnLeuGlulSerLysLeuLysProAspLeuLe 124  
Db 364 CCTAACTTAGAGGAATCTCTAAATTAACCGACTTAATT 405  
RESULT 10  
US-08-781-986A-238  
; Sequence 238, Application US/08781986A  
; Patent No. 6737248  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5255  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/781.986A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Benson, Bob  
; REGISTRATION NUMBER: 30,446  
; REFERENCE/DOCKET NUMBER: PB248PP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 238:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3775 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-781-986A-238  
Alignment Scores:  
Pred. No.: 0.000295 Length: 3775  
Score: 14.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.18% Indels: 0  
DB: 4 Gaps: 0  
US-10-724-972A-6352 (1-335) x US-08-781-986A-238 (1-3775)  
QY 111 ProAsnLeuGlulSerLysLeuLysProAspLeuLe 124  
Db 364 CCTAACTTAGAGGAATCTCTAAATTAACCGACTTAATT 405

PRIOR APPLICATION NUMBER: US 60/080,296  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: US 60/084,674  
; PRIOR FILING DATE: 1998-05-07  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 15  
; LENGTH: 999  
; TYPE: DNA  
; ORGANISM: Staphylococcus aureus  
US-10-278-946-15  
Alignment Scores:  
Pred. No.: 8.28e-05 Length: 999  
Score: 14.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.18% Indels: 0  
DB: 4 Gaps: 0  
US-10-724-972A-6352 (1-335) x US-10-278-946-15 (1-999)  
QY 111 ProAsnLeuGlulSerLysLeuLysProAspLeuLe 124  
Db 322 CCTAACTTAGAGGAATCTCTAAATTAACCGACTTAATT 363  
RESULT 9  
US-08-956-171E-238  
; Sequence 238, Application US/08956171E  
; Patent No 6593114  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; Gil H. Choi  
; Patrick S. Dillon  
; Craig A. Rosen  
; Steven C. Barash  
; Michael R. Fannon  
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5256  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/956,171E  
; FILING DATE: 20-Oct-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/009,861  
; FILING DATE: January 5, 1996  
; APPLICATION NUMBER: 08/781,986  
; FILING DATE: January 3, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mark J. Hyman  
; REGISTRATION NUMBER: 46,789  
; REFERENCE/DOCKET NUMBER: PB248P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (240) 314-1224  
; TELEFAX: (301) 309-8439  
; INFORMATION FOR SEQ ID NO: 238:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3775 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear

QY	169	GlyLysGluGluGluGlyLysArg	177
Db	7073	GGCAAGGAGGAGGAGGAAAGA	7047
RESULT 13			
US-09-949-016-14516/c			
; Sequence 14516, Application US/09949016			
; Patent No. 6812339			
; GENERAL INFORMATION:			
; APPLICANT: VENTER, J. Craig et al.			
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF			
; FILE REFERENCE: CL001307			
; CURRENT APPLICATION NUMBER: US/09/949,016			
; CURRENT FILING DATE: 2000-04-14			
; PRIOR APPLICATION NUMBER: 60/241,755			
; PRIOR FILING DATE: 2000-10-20			
; PRIOR APPLICATION NUMBER: 60/237,768			
; PRIOR FILING DATE: 2000-10-03			
; PRIOR APPLICATION NUMBER: 60/231,498			
; PRIOR FILING DATE: 2000-09-08			
; NUMBER OF SEQ ID NOS: 207012			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 14516			
; LENGTH: 203475			
; TYPE: DNA			
; ORGANISM: Human			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (1)..(203475)			
; OTHER INFORMATION: n = A,T,C or G			
US-09-949-016-14516			
Alignment Scores:			
Pred. No.:	1.82e+03	Length:	203475
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.6%	Indels:	0
DB:	4	Gaps:	0
US-10-724-972A-6352 (1-335) x US-09-949-016-14516 (1-203475)			
QY	32	SerRanSerLysGluSerSerLys	40
Db	67872	AGCAACTCTCCAGGAGAGCTCTAAG	67846
RESULT 14			
US-09-949-016-14517/c			
; Sequence 14517, Application US/09949016			
; Patent No. 6812339			
; GENERAL INFORMATION:			
; APPLICANT: VENTER, J. Craig et al.			
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF			
; FILE REFERENCE: CL001307			
; CURRENT APPLICATION NUMBER: US/09/949,016			
; CURRENT FILING DATE: 2000-04-14			
; PRIOR APPLICATION NUMBER: 60/241,755			
; PRIOR FILING DATE: 2000-10-20			
; PRIOR APPLICATION NUMBER: 60/237,768			
; PRIOR FILING DATE: 2000-10-03			
; PRIOR APPLICATION NUMBER: 60/231,498			
; PRIOR FILING DATE: 2000-09-08			
; NUMBER OF SEQ ID NOS: 207012			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 14517			
; LENGTH: 203475			
; TYPE: DNA			
; ORGANISM: Human			
; FEATURE:			
; NAME/KEY: misc_feature			

```
; LOCATION: (1)...(203475)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14517

Alignment Scores:
Pred. No.: 1.82e+03 203475
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.69%
DB: 4
Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-949-016-14517 (1-203475)
Qy 32 SerAenSerLysGluSerLys 40
Db 67872 AGCAACTCTCCAGGAGAGCTCTTAAG 67846

RESULT 15
US-09-949-016-14518/c
; Sequence 14518, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14518
; LENGTH: 203475
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)...(203475)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14518

Alignment Scores:
Pred. No.: 1.82e+03 203475
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.69%
DB: 4
Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-949-016-14518 (1-203475)
Qy 32 SerAenSerLysGluSerLys 40
Db 67872 AGCAACTCTCCAGGAGAGCTCTTAAG 67846

RESULT 16
US-09-949-016-14519/c
; Sequence 14519, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14519
; LENGTH: 203475
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)...(203475)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17226

Alignment Scores:
Pred. No.: 1.82e+03 203475
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.69%
DB: 4
Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-949-016-17226 (1-203475)
Qy 32 SerAenSerLysGluSerLys 40
Db 67872 AGCAACTCTCCAGGAGAGCTCTTAAG 67846

RESULT 17
US-09-949-016-17226/c
; Sequence 17226, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17226
; LENGTH: 203475
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)...(203475)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17226

Alignment Scores:
Pred. No.: 1.82e+03 203475
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.69%
DB: 4
Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-949-016-17226 (1-203475)
Qy 32 SerAenSerLysGluSerLys 40
Db 67872 AGCAACTCTCCAGGAGAGCTCTTAAG 67846
```

```
; LOCATION: (1)...(203475)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14517

Alignment Scores:
Pred. No.: 1.82e+03 203475
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.69%
DB: 4
Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-949-016-14517 (1-203475)
Qy 32 SerAenSerLysGluSerLys 40
Db 67872 AGCAACTCTCCAGGAGAGCTCTTAAG 67846

RESULT 15
US-09-949-016-14518/c
; Sequence 14518, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14518
; LENGTH: 203475
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)...(203475)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14518

Alignment Scores:
Pred. No.: 1.82e+03 203475
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.69%
DB: 4
Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-949-016-14518 (1-203475)
Qy 32 SerAenSerLysGluSerLys 40
Db 67872 AGCAACTCTCCAGGAGAGCTCTTAAG 67846

RESULT 16
US-09-949-016-14519/c
; Sequence 14519, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14519
; LENGTH: 203475
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)...(203475)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17226

Alignment Scores:
Pred. No.: 1.82e+03 203475
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.69%
DB: 4
Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-949-016-17226 (1-203475)
Qy 32 SerAenSerLysGluSerLys 40
Db 67872 AGCAACTCTCCAGGAGAGCTCTTAAG 67846

RESULT 17
US-09-949-016-17226/c
; Sequence 17226, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17226
; LENGTH: 203475
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)...(203475)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17226

Alignment Scores:
Pred. No.: 1.82e+03 203475
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.69%
DB: 4
Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-949-016-17226 (1-203475)
Qy 32 SerAenSerLysGluSerLys 40
Db 67872 AGCAACTCTCCAGGAGAGCTCTTAAG 67846
```

Alignment Scores: 1.82e+03 Length: 203475  
Pred. No.: 9.00 Matches: 9  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 2.69% Gaps: 0  
DB: 4

US-10-724-972A-6352 (1-335) x US-09-949-016-17228 (1-203475)

QY 32 SerAnSerLysGluSerLys 40  
Db 67872 AGCAACTCTCCAAGGAGGCTCTAAG 67846

RESULT 20  
US-09-949-016-17229/c  
; Sequence 17229, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17229  
; LENGTH: 203475  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(203475)  
; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-17227

Alignment Scores: 1.82e+03 Length: 203475  
Pred. No.: 9.00 Matches: 9  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 2.69% Gaps: 0  
DB: 4

US-10-724-972A-6352 (1-335) x US-09-949-016-17227 (1-203475)

QY 32 SerAnSerLysGluSerLys 40  
Db 67872 AGCAACTCTCCAAGGAGGCTCTAAG 67846

RESULT 19  
US-09-949-016-17228/c  
; Sequence 17228, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17228  
; LENGTH: 203475  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(203475)  
; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-17228

Alignment Scores: 1.82e+03 Length: 203475  
Pred. No.: 9.00 Matches: 9  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 2.69% Gaps: 0  
DB: 4

US-10-724-972A-6352 (1-335) x US-09-949-016-17229 (1-203475)

QY 32 SerAnSerLysGluSerLys 40  
Db 67872 AGCAACTCTCCAAGGAGGCTCTAAG 67846

RESULT 21  
US-09-513-999C-32507/c  
; Sequence 32507, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681

Alignment Scores: 1.82e+03 Length: 203475  
Pred. No.: 9.00 Matches: 9  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 2.69% Gaps: 0  
DB: 4

US-10-724-972A-6352 (1-335) x US-09-949-016-17227 (1-203475)

QY 32 SerAnSerLysGluSerLys 40  
Db 67872 AGCAACTCTCCAAGGAGGCTCTAAG 67846

RESULT 19  
US-09-949-016-17228/c  
; Sequence 17228, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17227  
; LENGTH: 203475  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(203475)  
; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-17227

Alignment Scores: 1.82e+03 Length: 203475  
Pred. No.: 9.00 Matches: 9  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 2.69% Gaps: 0  
DB: 4

US-10-724-972A-6352 (1-335) x US-09-949-016-17227 (1-203475)

QY 32 SerAnSerLysGluSerLys 40  
Db 67872 AGCAACTCTCCAAGGAGGCTCTAAG 67846

RESULT 19  
US-09-949-016-17228/c  
; Sequence 17228, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17228  
; LENGTH: 203475  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(203475)  
; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-17228



```
; SOFTWARE: Patent.pm
; SEQ ID NO 32507
; LENGTH: 181
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 16
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 130
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 131
; OTHER INFORMATION: n=a, g, c or t
; US-09-513-999C-32507

Alignment Scores:
Pred. No.: 23.6 Length: 181
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.39% Indels: 0
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-513-999C-32507 (1-181)
QY 30 SerSerSerAsnSerSerLysGlu 37
Db 163 TCTCTCTCNACTCTTCCAAAGAA 140

RESULT 22
US-09-313-294A-5824
; Sequence 5824, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Ialugudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 5824
; LENGTH: 305
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700350739H1
; US-09-313-294A-5824

Alignment Scores:
Pred. No.: 38.8 Length: 305
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.39% Indels: 0
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-313-294A-5824 (1-305)
QY 306 TrpAlaArgSerArgGlyLeuIle 313
Db 243 TGGCGCGGTTCAAGAGGTCTTATC 266

RESULT 23
US-09-513-999C-2950
; Sequence 2950, Application US/09513999C
; Patent No. 6783961
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59 US2 REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 2950
; LENGTH: 352
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 111...350
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 119
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 121
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 189
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 188
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 260
; OTHER INFORMATION: v=a or c or g
; NAME/KEY: UNSURE
; LOCATION: 4
; OTHER INFORMATION: Xaa=Arg or Thr
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 27
; OTHER INFORMATION: Xaa=Asp or His
; NAME/KEY: UNSURE
; LOCATION: 50
; OTHER INFORMATION: Xaa=Asp or Glu
; US-09-513-999C-2950

Alignment Scores:
Pred. No.: 44.5 Length: 352
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.39% Indels: 0
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-513-999C-2950 (1-352)
QY 28 AsnAsnSerSerSerAsnSerSer 35
Db 148 AACACAGCAGCAGCAGCAGCAGC 171

RESULT 24
US-09-513-999C-29145
; Sequence 29145, Application US/09513999C
; Patent No. 6783961
```

GENERAL INFORMATION:  
APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Wiegand, Roger C.  
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
FILE REFERENCE: 38-10(15849)B  
CURRENT APPLICATION NUMBER: US/09/902,540  
CURRENT FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: 60/217,883  
PRIOR FILING DATE: 2000-07-10  
NUMBER OF SEQ ID NOS: 16825  
SEQ ID NO 2334  
LENGTH: 564  
TYPE: DNA  
ORGANISM: Myxococcus xanthus  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(564)  
OTHER INFORMATION: unsure at all n locations  
US-09-902-540-2334

Alignment Scores:  
Pred. No.: 69.8 Length: 564  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.39% Indels: 0  
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-902-540-2334 (1-564)

Qy 306 TTPAlaArgSerArgGlyLeuIle 313  
Db 502 TGGCGCGCTCACGGCGCTCATC 525

RESULT 27  
US-09-949-016-39108  
Sequence 39108, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 39108  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-39108

Alignment Scores:  
Pred. No.: 74.2 Length: 601  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.39% Indels: 0  
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-949-016-39108 (1-601)

Qy 171 GluGluGlyLysLysArgLeu 178

GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Duclert, A.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
FILE REFERENCE: 59.US2.REG  
CURRENT APPLICATION NUMBER: US/09/513,999C  
CURRENT FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/122,487  
PRIOR FILING DATE: 1999-02-26  
NUMBER OF SEQ ID NOS: 36681  
SOFTWARE: Patent.pm  
SEQ ID NO 29145  
LENGTH: 378  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 271  
OTHER INFORMATION: m=a or c  
US-09-513-999C-29145

Alignment Scores:  
Pred. No.: 47.6 Length: 378  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.39% Indels: 0  
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-513-999C-29145 (1-378)

Qy 93 LysProLeuArgApLysIleGly 100  
Db 178 AAGCGCTAGAGATAAATAGGA 201

RESULT 25  
US-09-328-352-884  
Sequence 884, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Bratton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 884  
LENGTH: 390  
TYPE: DNA  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-884

Alignment Scores:  
Pred. No.: 49.1 Length: 390  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.39% Indels: 0  
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-328-352-884 (1-390)

Qy 238 LysGlyLeuSerLysTyrLeuLys 245  
Db 349 AAAGGCTTGTCTAAATATTAAAA 372

RESULT 26  
US-09-902-540-2334  
Sequence 2334, Application US/09902540  
Patent No. 6833447

DB 69 GAAGAAAGCGGAAAAAAGACTC 92

RESULT 28  
US-09-949-016-60203/c  
; Sequence 60203, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 60203  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-60203

Alignment Scores:
Pred. No.: 74.2 Length: 601
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.39% Indels: 0
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-949-016-60203 (1-601)

Qy 17 LeuPheValLeulleAlaThrAla 24  
|||||  
Db 145 CTCCTTGTCTCATTTGCTACTGCC 122

RESULT 29  
US-09-949-016-60204/c  
; Sequence 60204, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 60204  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-60204

Alignment Scores:
Pred. No.: 74.2 Length: 601
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.39% Indels: 0
DB: 4 Gaps: 0

Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.39% Indels: 0  
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-949-016-149234 (1-601)

QY 204 ValAlaAlaLysSerGlyLeuLeu 211  
DB 503 GTGGCTGTCACAAAGTGGGCTCTC 526

RESULT 32  
US-09-134-001C-909  
; Sequence 909, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 909  
; LENGTH: 789  
; TYPE: DNA  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-909

Alignment Scores:  
Pred. No.: 96.2 Length: 789  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.39% Indels: 0  
DB: 3 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-134-001C-909 (1-789)

QY 24 AlaAlaCysGlyAsnAsnSerSer 31  
DB 46 GCAGCCTGTGGCAACAATTCATCT 69

RESULT 33  
US-09-710-279-1567  
; Sequence 1567, Application US/09710279  
; Patent No. 6703492  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PUS480US  
; CURRENT APPLICATION NUMBER: US/09/710,279  
; CURRENT FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1567  
; LENGTH: 789  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-09-710-279-1567

Alignment Scores:  
Pred. No.: 96.2 Length: 789

Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.39% Indels: 0  
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-710-279-1567 (1-789)

QY 24 AlaAlaCysGlyAsnAsnSerSer 31  
DB 46 GCAGCCTGTGGCAACAATTCATCT 69

RESULT 34  
US-09-543-681A-1850  
; Sequence 1850, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 1850  
; LENGTH: 843  
; TYPE: DNA  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-1850

Alignment Scores:  
Pred. No.: 103 Length: 843  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.39% Indels: 0  
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-543-681A-1850 (1-843)

QY 119 LeuLysProAspLeuIleIleAla 126  
DB 256 TTAAGACCTGATTAAATTAATTCATCT 279

RESULT 35  
US-09-270-767-14140  
; Sequence 14140, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 14140  
; LENGTH: 886  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-14140

Alignment Scores:  
Pred. No.: 108 Length: 886  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.39% Indels: 0  
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-270-767-14140 (1-886)

OTHER INFORMATION: 12-515-394.mis1, potential  
NAME/KEY: misc binding  
LOCATION: 502..521  
OTHER INFORMATION: 12-515-394.mis2, potential complement  
NAME/KEY: primer\_bind  
LOCATION: 107..126  
OTHER INFORMATION: upstream amplification primer  
NAME/KEY: primer\_bind  
LOCATION: 537..557  
OTHER INFORMATION: downstream amplification primer, complement  
NAME/KEY: misc binding  
LOCATION: 489..513  
OTHER INFORMATION: 12-515-394 potential probe  
US-09-641-638-192

Alignment Scores:  
Pred. No.: 121 Length: 1001  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.39% Indels: 0  
DB: 3 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-641-638-192 (1-1001)

QY 29 AsnSerSerSerSerSerLys 36  
|||||  
Db 923 AATTCTAGCTCTACAGCTCTAAA 946  
|||||

RESULT 38  
US-09-641-638-193  
Sequence 193, Application US/09641638  
Patent No. 6432648  
GENERAL INFORMATION:  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Bougueleret, Lydie  
APPLICANT: Chumakov, Ilya  
APPLICANT: Cohen, Annick  
TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING  
FILE REFERENCE: GENSET.051CP1  
CURRENT APPLICATION NUMBER: US/09/641,638  
PRIOR FILING DATE: 2000-08-16  
PRIOR APPLICATION NUMBER: US 09/502,330  
PRIOR FILING DATE: 2000-02-11  
PRIOR APPLICATION NUMBER: US 60/133,200  
PRIOR FILING DATE: 1999-05-07  
PRIOR APPLICATION NUMBER: US 09/275,267  
PRIOR FILING DATE: 1999-03-23  
PRIOR APPLICATION NUMBER: US 60/119,917  
PRIOR FILING DATE: 1999-02-12  
NUMBER OF SEQ ID NOS: 1304  
SOFTWARE: Patent.pm  
SEQ ID NO 193  
LENGTH: 1001  
TYPE: DNA  
ORGANISM: Homo Sapiens  
FEATURE:  
NAME/KEY: allele  
LOCATION: 501  
OTHER INFORMATION: 12-516-97 : polymorphic base C or T  
NAME/KEY: misc binding  
LOCATION: 481..500  
OTHER INFORMATION: 12-516-97.mis1, potential  
NAME/KEY: misc binding  
LOCATION: 502..520  
OTHER INFORMATION: 12-516-97.mis2, complement  
NAME/KEY: primer\_bind  
LOCATION: 405..424  
OTHER INFORMATION: upstream amplification primer  
NAME/KEY: primer\_bind  
LOCATION: 891..911  
OTHER INFORMATION: downstream amplification primer, complement

QY 28 AsnSerSerSerSerSerSer 35  
|||||  
Db 726 ACAAATAGCAGCAGCAACAGCAGC 749  
|||||

RESULT 36  
US-09-489-039A-2246  
Sequence 2246, Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:  
APPLICANT: Gary Breton et. al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 2709.2004001  
CURRENT APPLICATION NUMBER: US/09/489,039A  
CURRENT FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: US 60/117,747  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 2246  
LENGTH: 981  
TYPE: DNA  
ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-2246

Alignment Scores:  
Pred. No.: 119 Length: 981  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.39% Indels: 0  
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-489-039A-2246 (1-981)

QY 203 AlaValAlaAlaLysSerGlyLeu 210  
|||||  
Db 678 GCCGTGGCTGCANAAAGCGACTG 701  
|||||

RESULT 37  
US-09-641-638-192  
Sequence 192, Application US/09641638  
Patent No. 6432648  
GENERAL INFORMATION:  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Bougueleret, Lydie  
APPLICANT: Chumakov, Ilya  
APPLICANT: Cohen, Annick  
TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING  
FILE REFERENCE: GENSET.051CP1  
CURRENT APPLICATION NUMBER: US/09/641,638  
CURRENT FILING DATE: 2000-08-16  
PRIOR APPLICATION NUMBER: US 09/502,330  
PRIOR FILING DATE: 2000-02-11  
PRIOR APPLICATION NUMBER: US 60/133,200  
PRIOR FILING DATE: 1999-05-07  
PRIOR APPLICATION NUMBER: US 09/275,267  
PRIOR FILING DATE: 1999-03-23  
PRIOR APPLICATION NUMBER: US 60/119,917  
PRIOR FILING DATE: 1999-02-12  
NUMBER OF SEQ ID NOS: 1304  
SOFTWARE: Patent.pm  
SEQ ID NO 192  
LENGTH: 1001  
TYPE: DNA  
ORGANISM: Homo Sapiens  
FEATURE:  
NAME/KEY: allele  
LOCATION: 501  
OTHER INFORMATION: 12-515-394 : polymorphic base A or T  
NAME/KEY: misc binding  
LOCATION: 481..500

US-10-170-097-192

Alignment Scores:

Pred. No.:	121	Length:	1001
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.39%	Indels:	0
DB:	4	Gaps:	0

US-10-724-972A-6352 (1-335) x US-10-170-097-192 (1-1001)

QY 29 AnSersSerAnSersSerLys 36  
|||||

Db 923 AATTCTAGCTTAACAGCTCTAAA 946  
|||||

RESULT 40

US-10-170-097-193  
; Sequence 193, Application US/10170097  
; Patent No. 6794143  
; GENERAL INFORMATION:  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Cohen, Annick  
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING  
; FILE REFERENCE: GEN-T114XC2D1  
; CURRENT APPLICATION NUMBER: US/10/170,097  
; CURRENT FILING DATE: 2002-06-10  
; PRIOR APPLICATION NUMBER: US 09/641,638  
; PRIOR FILING DATE: 2000-08-16  
; PRIOR APPLICATION NUMBER: US 09/502,330  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: US 60/133,200  
; PRIOR FILING DATE: 1999-05-07  
; PRIOR APPLICATION NUMBER: US 09/275,267  
; PRIOR FILING DATE: 1999-03-23  
; PRIOR APPLICATION NUMBER: US 60/119,917  
; PRIOR FILING DATE: 1999-02-12  
; NUMBER OF SEQ ID NOS: 1304  
; SOFTWARE: Patent.pm  
; SEQ ID NO 193  
; LENGTH: 1001  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 501  
; OTHER INFORMATION: 12-516-97 : polymorphic base C or T  
; FEATURE:  
; NAME/KEY: misc binding  
; LOCATION: 481..500  
; OTHER INFORMATION: 12-516-97.mis1, potential  
; FEATURE:  
; NAME/KEY: misc binding  
; LOCATION: 502..520  
; OTHER INFORMATION: 12-516-97.mis2, complement  
; FEATURE:  
; NAME/KEY: primer bind  
; LOCATION: 405..424  
; OTHER INFORMATION: upstream amplification primer  
; FEATURE:  
; NAME/KEY: primer bind  
; LOCATION: 891..911  
; OTHER INFORMATION: downstream amplification primer, complement  
; FEATURE:  
; NAME/KEY: misc binding  
; LOCATION: 489..513  
; OTHER INFORMATION: 12-516-97 potential probe  
US-10-170-097-193

Alignment Scores:

US-10-724-972A-6352 (1-335) x US-09-641-638-193 (1-1001)

QY 29 AnSersSerAnSersSerLys 36  
|||||

Db 717 AATTCTAGCTTAACAGCTCTAAA 740  
|||||

RESULT 39

US-10-170-097-192  
; Sequence 192, Application US/10170097  
; Patent No. 6794143  
; GENERAL INFORMATION:  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Cohen, Annick  
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING  
; FILE REFERENCE: GEN-T114XC2D1  
; CURRENT APPLICATION NUMBER: US/10/170,097  
; CURRENT FILING DATE: 2002-06-10  
; PRIOR APPLICATION NUMBER: US 09/641,638  
; PRIOR FILING DATE: 2000-08-16  
; PRIOR APPLICATION NUMBER: US 09/502,330  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: US 60/133,200  
; PRIOR FILING DATE: 1999-05-07  
; PRIOR APPLICATION NUMBER: US 09/275,267  
; PRIOR FILING DATE: 1999-03-23  
; PRIOR APPLICATION NUMBER: US 60/119,917  
; PRIOR FILING DATE: 1999-02-12  
; NUMBER OF SEQ ID NOS: 1304  
; SOFTWARE: Patent.pm  
; SEQ ID NO 192  
; LENGTH: 1001  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 501  
; OTHER INFORMATION: 12-515-394 : polymorphic base A or T  
; FEATURE:  
; NAME/KEY: misc binding  
; LOCATION: 481..500  
; OTHER INFORMATION: 12-515-394.mis1, potential  
; FEATURE:  
; NAME/KEY: misc binding  
; LOCATION: 502..521  
; OTHER INFORMATION: 12-515-394.mis2, potential complement  
; FEATURE:  
; NAME/KEY: primer bind  
; LOCATION: 107..126  
; OTHER INFORMATION: upstream amplification primer  
; FEATURE:  
; NAME/KEY: primer bind  
; LOCATION: 537..557  
; OTHER INFORMATION: downstream amplification primer, complement  
; FEATURE:  
; NAME/KEY: misc binding  
; LOCATION: 489..513  
; OTHER INFORMATION: 12-515-394 potential probe  
US-10-724-972A-6352 (1-335) x US-09-641-638-193 (1-1001)

Alignment Scores:

Pred. No.:	121	Length:	1001
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.39%	Indels:	0
DB:	3	Gaps:	0

1000000 2005

Pred. No.:	121	Length:	1001
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.39%	Indels:	0
DB:	4	Gaps:	0

US-10-724-972A-6352 (1-335) x US-10-170-097-193 (1-1001)

Qy	29	AsnSerSerAenSerSerLys	36
Db	717	AATTCTAGCTTAACAGCTCTAAA	740

Search completed: November 10, 2005, 03:07:59  
Job time : 664 secs

This page Blank (uspdc)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 9, 2005, 19:43:37 ; Search time 5171 Seconds  
(without alignments)  
3139.141 Million cell updates/sec

Title: US-10-724-972A-6352

Perfect score: 1690

Sequence: 1 GVESVRLKILSVIGLLFVL.....EEMAKELVELSKDKKDKK 335

Scoring table:  
BLOSUM62 Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO.spool.h/US10724972/runat\_03112005.141859.29146/app\_query.fasta\_1.519  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10724972 @cgn 1.1 3731 @runat\_03112005.141859.29146 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT\_DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.ats.\*
- 12: gb.sv.\*
- 13: gb.un.\*
- 14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
C 1	1690	100.0 300698 1	AE016750 Staphylococcus epidermidis ATCC 12228
C 2	1233	73.0 307750 1	AP003136 Staphylococcus epidermidis ATCC 12228
C 3	1233	73.0 347650 1	AP003364 Staphylococcus epidermidis ATCC 12228
C 4	1231	72.8 110000 1	Continuation (24 o

ALIGNMENTS

RESULT 1	AE016750/c	AE016750	300698 bp	DNA	linear	BCT 01-JAN-2003
LOCUS	Staphylococcus epidermidis ATCC 12228	complete genome.				section 7 of 9 of the
DEFINITION	AE016750 AE015929	AE016750.1	GI:27316220			
ACCESSION	AE016750 AE015929					
VERSION	AE016750.1					
KEYWORDS	Staphylococcus epidermidis ATCC 12228					
SOURCE	Staphylococcus epidermidis ATCC 12228					
ORGANISM	Bacteria; Firmicutes; Bacillales; Staphylococcus.					
REFERENCE	1 (bases 1 to 300698)					
AUTHORS	Zhang, Y., Ren, S., Li, H., Fu, G., Lu, L., Jia, J., Tu, Y., Qin, Z., Chen, Z. and Wen, Y.					
TITLE	Direct Submission					
JOURNAL	Submitted (05-NOV-2002) Chinese National Human Genome Center at Shanghai, 250 Bi Bo Road, Shanghai 201203, China					
FEATURES	Location/Qualifiers					
source	1. 300698					
	/organism="Staphylococcus epidermidis ATCC 12228"					
	/mol_type="genomic DNA"					
	/strain="ATCC 12228"					
	/db_xref="ATCC:12228"					
	/db_xref="taxon:176280"					
	complement(114..186)					
	/product="tRNA-Lys"					
	complement(190..264)					
	/product="tRNA-Gln"					

2



```
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AA005405.1"
/db_xref="GI:27316230"
/transl_table="MFVYTRITVKKGAKQMANFTKGGPIESLKGFEQVWQIDK
DDYSEDYVNSWETEEDFKNNVNSDFKQAHKNTGKSEDSPIKSEIVKSNVLSLN
RR"
complement(14682..16055)
/locus_tag="SE1765"
complement(14682..16055)
/locus_tag="SE1765"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AA005406.1"
/db_xref="GI:27316231"
/transl_table="MNGPKVDIVRGRIKFTMSLGIILFLVPVQDGKQOITLP
IAPLAGLKDMLGIMIPILITITVSGIILCTIYKNKLNPOGLMSAFNKGW
LLEKLVAFIPFISWTLFNLGEMIKSEDTGGLVFSLLPLTAVLFALPLPLMEYG
LBUKGLPIFRPPLFTLFRSTVDNLSAFIGDGTGVLITSRQEGEYSRREAV
ISTFSVSTFIAVTAETIRMQDFYFYLTVVISCLIAAMIPRIPLKNTPDSEA
KEYSEARNEQLPEGTALKYGFDLATEVIGIKSPKPEFLISGFKTVDMFWILFPV
MSGTATTIANTVPFEIIGKFPVPLELLOIPEAHEASQTILIGFADMFDPSILIE
GQNDVTRFVIGALSISQLVILSEVGVILGSKIPVSISKLFMIFLIRTIITPLIAL
LAHLFIG"
complement(16352..17317)
/locus_tag="SE1766"
```

gene

CDS

gene

## Alignment Scores:

```
Pred. No.: 2,67e-115 Length: 300698
Score: 1690.00 Matches: 335
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0
```

US-10-724-972A-6352 (1-335) x AE016750 (1-300698)

```
Qy 1 GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuPheValLeu 20
Db 19364 GGAGTGAATCAGTGAGAGGTTTAAATAATTTTAAGTGTAATGGCTTATTGTTGTTTA 19305
Qy 21 IleAlaThrAlaAlaCysGlyAsnAsnSerSerAsnSerSerLysGluSerLys 40
Db 19304 ATTGCAACTGCAGCATGTGCAATAATATAGTTCAAGTAACCTCAAGTAAGAGTCATCAAAA 19245
Qy 41 AspGlyValGluIleLysHisGluGluGlyThrLysValProLysHisProLysArg 60
Db 19244 GATGGAGTTGAAATCAAGCAGCAAGAGGTACTAGCAAGTACCTAAACACCCCTAAAGCT 19185
Qy 61 ValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProVal 80
Db 19184 GTTGTGTTCTTGAGTATTCATTGTTGATGCGTTAGTTGCTTTAGATGTTTAAACCTGTT 19125
Qy 81 GlyIleAlaAspAspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGly 100
Db 19124 GGGATAGCGGATGATAACAAAAAATTCGTATTATTAACCATTAAGAGATAAAATTTGA 19065
Qy 101 LysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLys 120
Db 19064 AAATACACTCTGTAGGAACACGTAAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAA 19005
Qy 121 ProAspLeuIleIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLys 140
Db 19004 CCAGATTTAATTTGCTGATAATAATAGACACAAAGGTATTTATAAGACTTAAATAAA 18945
Qy 141 IleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAla 160
Db 18944 ATTGCTCTCAGATTGAACCTGAAAGTTTCAGTGGAGATTAATGAAAAATTTGATGCT 18885
Qy 161 PheLysThrIleSerLysAlaLeuGlyLysGluGluGlyLysArgLeuGlu 180
```

```
Db 18884 TTTAAACAATTTTCAAAGCTTTAGGTAAAGAGAAGAAAGGTAATAAACCGCTTGAAGAA 18825
Qy 181 HisAspLysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysVal 200
Db 18824 CACGATAAGAAAAATTTGAAGAATAATAAAAAAGAAATAACTATGCTATAAAATCAAAAGGTA 18765
Qy 201 LeuProAlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGly 220
Db 18764 TTGCTGCAGTAGCTGCTAAATCAGGTTTGTCTCATCAAGCAACTCTTATGTTGGT 18705
Qy 221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu 240
Db 18704 CAATTCCTAAGTCAACTAGGTTTAAAGAGCAATTAAGTATGATGTTACTAAGGTTTA 18645
Qy 241 SerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsn 260
Db 18644 AGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAACTTTATCTCAAGTCAAT 18585
Qy 261 ProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGlu 280
Db 18584 CCTGACCGTATGTTCAATATGACAAACAAAGCAAGTTCTTAACGAACTTCTCACTAAAGAA 18525
Qy 281 LeuGluLysAspProValTyrLysLysLeuAsnAlaValLysAsnGlnArgValAspIle 300
Db 18524 CTAGAAAAAGATCCTGTATGGAAGAAATTAACCGCTGTGAATAATCAACGTGTTGATATT 18465
Qy 301 LeuAspArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLys 320
Db 18464 TTAGACCGTACTTATGGGCAAGATCACGTGTTTAAATTTCTTCAAGAAATGCGCAAAA 18405
Qy 321 GluLeuValGluLeuSerLysLysAspSerLysLysAspAsnLys 335
Db 18404 GAACTTGTTGAATATTCTAAGAAAGATAGTAAAAAAAGATAATAAG 18360
RESULT 2
AP003136/c 307750 bp DNA linear BCT 11-JAN-2003
LOCUS Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
DEFINITION AP003136 BA000018
ACCESSION AP003136.2 GI:14349228
VERSION 1
KEYWORDS Staphylococcus aureus subsp. aureus N315
SOURCE Staphylococcus aureus subsp. aureus N315
ORGANISM Staphylococcus aureus subsp. aureus N315
REFERENCE 1
AUTHORS Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Iian,J., Ito,T., Kanamori,M.,
Matsumaru,H., Maruyama,A., Murakami,H., Hoshino,A.,
Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
Kanehisa,M., Yanahita,A., Oshima,K., Furuya,K., Yoshino,C.,
Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
Whole genome sequencing of methicillin-resistant Staphylococcus
aureus
JOURNAL Lancet 357 (9264), 1225-1240 (2001)
MEDLINE 21311952
PUBMED 11418146
REFERENCE 2 (bases 1 to 307750)
AUTHORS Director-General, Biotechnology Center, Aoki,K., Oguchi,A.,
Hosoyama,A., Nagai,Y., Kuroda,M., Hiramatsu,K. and Kikuchi,H.
Direct Submission
Submitted (30-JAN-2001) Director-General, Biotechnology Center,
National Institute of Technology and Evaluation, Biotechnology
Center, 2Chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
(E-mail:bioelite.go.jp, URL:http://www.bio.nite.go.jp/,
Tel:81-3-3481-1933, Fax:81-3-3481-8424)
COMMENT On Jun 12, 2001 this sequence version replaced gi:13701842.
FEATURES Location/Qualifiers
source 1. 307750
/organism="Staphylococcus aureus subsp. aureus N315"
/mol_type="genomic DNA"
/strain="N315"
```

gene	/sub_species="aureus N315"		CDS	complement (4320. .4781)	
	/db_xref="taxon:158879"			/gene="SA1857"	
	297. .1907			/note="ORFID:SA1857"	
	/gene="SA1853"			/codon_start=1	
	297. .1907			/transl_table=11	
	/gene="SA1853"			/product="conserved hypothetical protein"	
	/note="ORFID:SA1853			/protein_id="BAB43139.1"	
	hypothetical protein, similar to DNA mismatch repair			/db_xref="GI:13701847"	
	protein MutS			/translation="MIKINNDEMNOFAMFLVEQLKSGDLILLNGDLGAGKTLTQFI	
	/codon_start=1			GKALGVRRITNSPTFNIISKYRGNLKLHMDCYRLSDSDLDGFDFEFQQAIVIE	
gene	/transl_table=11		CDS	WSQFIKDLPLATHLSINISTISENTRQIELFAQGHEYEIQEAIHEFAAH"	
	/protein_id="BAB43135.1"			5292. .6980	
	/db_xref="GI:13701843"			/gene="ilvD"	
	/translation="MSTNQFLIFVIAILLTSVIGIVGRYSRQRLKSMETLWQTI			5292. .6980	
	SPLEAFIRPNSHDFEYKLYKEFESHSLVDDKTWSDLNNAIFHKNNVNLTAIGEMK			/gene="ilvD"	
	LYACLRGMLSTNKSLLSPNDNAEPKNTYHLALIGKTYTFPDDITPQVRQNIL			/note="ORFID:SA1858"	
	FLCPLPVSIPAFINSQVGIILFLMSCLFNLSATLKRTEYEDDLKSIFYASNVLK			/codon_start=1	
	QYTKISKIKHAQPEVNFQKPRTAHLTSVLAENVDEDDIGAMVIKVLKLFMDLYLF			/transl_table=11	
	HSTQSKYTHMELKNCFDYLAEIDNHYALAMRTRLETCVTEQIIDDNDGIVSELT			/product="dihydroxy-acid dehydratase"	
	HPLIADAVANDFSLSQNLITGTSNASKSTFMKSIAINIILASAIQVTTASKFYOPG			/protein_id="BAB43140.1"	
gene	IVFTSMANADDVLSGDSYFMABLKSIIKRIVEIPDNQKIYCFIDEIFKGTWYTERIAS		CDS	/db_xref="GI:13701848"	
	ESVLSFHEKSNFVIAATHDIELAELLKQRYENHPFVNIENNIIHFDYIKRPGKAN			/translation="MRSMDIKKGQHQAPARSLHATGALKSPTDMNKPFAICNSYID	
	TRNAIELKITSPPAKIYERAKDNVSG"			IVPHVHLRELADIKAERAGALPFEPNTIGVDDGIAMGHIGMRYSLPSREIADA	
	complement (2222. .3247)			AETVINARELFDGVFIYPCNDKIIPGMILAAKRTNVPAIFCSGPGMKAGLSAHGKALT	
	/gene="SA1854"			SSMEFAVGAFKEGSIKEEFLDMEQNACPTCSGACMFTANSMCNMVEGLGALPYNG	
	complement (2222. .3247)			TALAUSDORREMIQAAQKLVENIKNDLKPRDITVREIDAIPALDMDMVGSTNTVLH	
	/gene="SA1854"			TLATANEAGIDYDLERINAIKRTPYLSKIAPSSYSMDHDEAGGVPAINELMKKD	
	/note="ORFID:SA1854			GTLHPDRITVTGKTLRENNEKEIKNFVDIHLDPADYDAQGLSLFGNIAKPGAVIK	
	hypothetical protein, similar to O-sialoglycoprotein			VGVDPDSITGTGAICFNCSDHDEAVEAIDNRTVRAGHVVIVREYEGKPGGPMPEMLAP	
	endopeptidase"			TSSIVGRGLGKDALITDGRFSGATRGIAVGHISPEAASGGPIALLIEGDEITLITL	
gene	/codon_start=1		CDS	RTLNVNQPEDVLARRRESLTPFKAKVTGYLTARYALTALVTSANTGGVMQVPEMLI"	
	/transl_table=11			7008. .8777	
	/protein_id="BAB43136.1"			/gene="ilvB"	
	/db_xref="GI:13701844"			7008. .8777	
	/translation="WTKDILILAVETSODETSVSVIKNRGDIILNTVLISQIESHKRFG			/gene="ilvB"	
	GVPEVASRHVEGTTTINRALVDADVSMDIDAIAVTEGPGIGLIGLIGNAAKAL			/note="ORFID:SA1859"	
	AFADKPLIPVHTAGHYANHIEEPLTFPIALIVSGHTELVYMKDHLSEFVIGET			/codon_start=1	
	RDDAVGEAYDKARTIGUNYPGQVQVDRLAEGEDTSPFRWLDKSDYSPFGLKS			/transl_table=11	
	AVINQLHQRQRQPIIEANVATSPQNSVSVVEVLTAKIACKESVQRLIIVAGGVASN			/product="acetolactate synthase large subunit"	
	KGLRSLADQCKVNDIQLTIPSKLCTDNAAMIGVAGSHLYQQRFADLALHNGSHNID			/protein_id="BAB43141.1"	
gene	LEEVSAAE"		CDS	/db_xref="GI:13701849"	
	complement (3240. .3704)			/translation="MSKTOHEVNQNIIDPLKMAESELEQLNEKTLNDMRSGSEVLVEA	
	/gene="SA1855"			LLKENVDYLFYGGGAVLPLDYDTFDGKIHKHILARHSQGAHAAEGVARVSGKTVGV	
	complement (3240. .3704)			VTSFGATNWTGITDAHCDLPLVFTFGVATPGIGKADAFQADILLMTSPITKQNY	
	/gene="SA1855"			QVKGVEDIPKIVHEAFVANSGRKGPVVIDFPKDMGVLATNVDLUCDEINIPGYEVTE	
	/note="ORFID:SA1855			PENKIDITFISLKEAKKPVLAGINQSKNQLTFQVNNKHQITPTVTLLGLGAVP	
	hypothetical protein, similar to ribosomal-protein-alanine			VEDTLFLGGMGHSYASNMALTECDLLINLGSRFDRLSKPAFAFAKIVHVDID	
	N-acetyltransferase"			PSEINKVHVDLGIADCKRFLCNDKNVETIEHSDWKHCQNNKHQKPFKLGEEDQ	
	/codon_start=1			VFCXPOQTIEYIKTNGEAIYTTDVGQHQMWAQFYPFKNHGQWVTSGGLTGWQKQW	
	/transl_table=11			PSSIGAKLANPDKTVCVGVGGVQFMQENALLPEYGLDKVILINNGTLGMWKQWQ	
gene	/protein_id="BAB43137.1"		CDS	DKFFNORFSSHVSFNGQDPFMKMAEAYGVKGFLLIDKPEQLBEQDAAAFAYQGPVLIEVR	
	/db_xref="GI:13701845"			1SPTEAVTPMVPSGKSNHEMGL"	
	/translation="WDQSKQELNIREMTKEDVPQVFDLERRSFNDSSWTIDAFYHEI			8777. .9031	
	EQNNPAKYPVLFEEQILIGLGLWIVDAQIITVAIDDNRYGYGLQMLLYGINYA			/gene="SA1860"	
	SHTCDVMSLEVRVNNKVAQVYENLFGYGGKRNKYGGEDAMVWVNLND"			8777. .9031	
	complement (3677. .4366)			/gene="SA1860"	
	/gene="SA1856"			/note="ORFID:SA1860	
	complement (3677. .4366)			hypothetical protein, similar to acetolactate synthase	
	/gene="SA1856"			small subunit"	
	/note="ORFID:SA1856			/codon_start=1	
gene	hypothetical protein, similar to glycoprotein		CDS	/transl_table=11	
	endopeptidase"			/protein_id="BAB43142.1"	
	/codon_start=1			/protein_id="BAB43140.1"	
	/transl_table=11			/db_xref="GI:13701850"	
	/protein_id="BAB43138.1"			/translation="MTRILKLOVADQVSTLNRTISAFVRLQYNIDTLVHTHSEQGIS"	
	/db_xref="GI:13701846"			NMEIQVDIQQDTSLLHLIKLKKQINVLTVCEYDLVDNEA"	
	/translation="MKNLRLQSLNSLLIDTSNQPLSVALMONDELVAELITDTSKONH			9168. .10172	
	SVQLMPAIGOLFQSKIAKQDLDAIIVAEFGSYTGIRGTVTAKTLAYALVDKLVGV			/gene="ilvC"	
	SSLKALAAITDHTDKLLVPVDFARQAVYTGIFQWQNGQLITLEDQYMSIEARLAF			9168. .10172	
	EDLNPFTYIGKDTIKLQDELQGEVIAQLPNASVMYQLIDKPTDVHTFTPKYHKLAAE			/gene="ilvC"	
gene	ERNWINSQNN"		CDS	/note="ORFID:SA1861"	
	complement (4320. .4781)			/codon_start=1	
	/gene="SA1857"			/transl_table=11	
				/product="alpha-keto-beta-hydroxylacil reductoisomerase"	

```

/protein_id="BAB43143.1"
/db_xref="GI:13701951"
/translations="MTVTYDDQDVKTALQGGKIAVVGSGQGHAAQNLIKNGDYDV
IGIRPGRSPDKAEDGDFVPAEAVKQADITMWLLPDEIOGDVYKNEISPNLEKHA
LAFAGRNTHFGVIOQPADVDVLPAKPGRGLHVRTPEGSAVPSIFGQQDASGA
RNTALSVAKGIGATRAGVLETFEKEETEDLFGEOALVCGGSKLIQSGSETLVEAGY
OPELAVFVLHEKMLILVDLMYGGENVRYSISNTAEFGDYVSGPRVITTDVKNENKA
VLTDIQNGFNSNFIEDNKGQKFKFYKLRERQHGHQIEKVGRELREMMPTIKSKTEK
"
gene
10202..11731
/genes="leuA"
10202..11731
/genes="leuA"
/notes="ORFID:SA1862"
/codon_start=1
/transl_table=11
/product="2-isopropylmalate synthase"

```

**Alignment Scores:**

Pred. No.:	6.27e-81	Length:	307750
Score:	1233.00	Matches:	238
Percent Similarity:	84.59%	Conservative:	42
Best Local Similarity:	71.90%	Mismatches:	51
Query Match:	72.96%	Indels:	1
DB:	1	Gaps:	0

US-10-724-972A-6352 (1-335) x AP003136 (1-307750)

QY	1	GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeu	20
Db	156662	CGAGTGGAAACG-ATCAGAGGCTTAAAAACTTTTAGTATATTGGATTATAGTTGCCTTA	156604
QY	21	IleAlaThrAlaAlaCysGlyAsnAsnSerSerAsnSerSerLysGluSerSerLys	40
Db	156603	CTTTTAGTTGCAGCTTGTGGTAATACGGATAATTCAGTAAAGATAAATCATCACTAA	156544
QY	41	AspGlyValGluIleLysHisGluGluGlyThrThrLysValProLysHisProLysArg	60
Db	156543	GATACTATTTTCGGTAAAGATGAAATGGTACAGTAAAGTACCTAAAGTCAAAACGT	156484
QY	61	ValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProVal	80
Db	156483	ATCGTTGTATTAGAGTACTCATTTGCAGATGCATTAGCAGATTAGACGTTAAACCGT	156424
QY	81	GlyIleAlaAspAspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGly	100
Db	156423	GGTATTGCTGATGATGGTAAGAAAAAACGTTATCATTAACACCTTAGAGAAAATTTGG	156364
QY	101	LysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluIleSerLysLeuLys	120
Db	156363	GATTATATCTCTGTAGGTACACGTAACAGCCAAACTTAGAGAAATTAGTAAATTA	156304
QY	121	ProAspLeuIleIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLys	140
Db	156303	CCGGATTTAATTATCGTGTATACGAGTGAACATAAAGGTATTAAATAAGAAATTAACAA	156244
QY	141	IleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAla	160
Db	156243	ATTGCACCAACATTATCATTTAAAGAGTTTTGTATGGAGACTACAAACAAATATTAA	156184
QY	161	PheLysThrIleSerLysAlaLeuGlyLysGluGluGluGlyLysLysArgLeuGlu	180
Db	156183	TTCAAAACAATTGCTAAAGCTTTAAATAAGAAAAAGAGGCGCAAAACGCTTGCTGAG	156124
QY	181	HisAspLysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysVal	200
Db	156123	CATGATAAATTATCAATAGTATAAGATGAATTAATTTGATAGAAATCAAAAAGTG	156064
QY	201	LeuProAlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGly	220
Db	156063	CTTCCAGCAGTAGTTGCTAAAGCTGTTTATTAGCACATCCAAACTATTCAATATGTTGGA	156004
QY	221	GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAspValThrLysGlyLeu	240

[illegible]

RESULT 3  
AP003364/C

LOCUS	AP003364	347650 bp	DNA	linear	BCT 17-APR-2004
DEFINITION	Staphylococcus aureus subsp. aureus Mu50 DNA, complete genome,				
ACCESSION	AP003364	BA000017			
VERSION	AP003364.3	GI:46395538			
KEYWORDS					
SOURCE	Staphylococcus aureus subsp. aureus Mu50				
ORGANISM	Staphylococcus aureus subsp. aureus Mu50				
	Bacteria; Firmicutes; Bacillales; Staphylococcus.				

REFERENCE AUTHORS	TITLE	JOURNAL MEDLINE PUBLISHED
Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I., Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Lian, J., Ito, T., Kanamori, M., Matsumaru, H., Maruyama, A., Murakami, H., Hoshoyama, T., Mizutani-Ui, Y., Takahashi, N.K., Sawano, T., Inoue, R., Kaito, C., Sekimizu, K., Hirakawa, H., Kikura, S., Goto, S., Yabuzaki, J., Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C., Shiba, T., Hattori, M., Ogasawara, N., Hayashi, H. and Hiramatsu, K.	Whole genome sequencing of methicillin-resistant <i>Staphylococcus aureus</i>	Lancet 357 (9264), 1225-1240 (2001) 21311952 11418146

2	11418146	Ohta, T., Hirakawa, H., Morikawa, K., Maruyama, A., Inose, Y., Yamashita, A., Oshima, K., Kuroda, M., Hattori, M., Hiramatsu, K., Kuhara, S. and Hayashi, H. Nucleotide Substitutions in <i>Staphylococcus aureus</i> Strains, Mu50, MuJ3, and N315 DNA RESEARCH 11, 51-56 (2004) 3 (bases 1 to 347650) Hattori, M., Yamashita, A., Oshima, K. and Shiba, T. Direct Submission Submitted (28-FEB-2001) Masahira Hattori, Kitasato Institute for Life Sciences, Kitasato University, Kitasato 1-15-1, Sagamihara, Kanagawa 228-8555, Japan (E-mail: hattori@genome.ils.kitasato-u.ac.jp, Tel: 81-42-778-8194, Fax: 81-42-778-8193)
FORMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		

COMMENT

April 16, 2004 this sequence version replaced g1:14247707. This work was done in collaboration with Toshiko Mutsaers, Kanamori, Hideo Hayashi (University of Tsukuba), Hideki Hirakawa, Satoru Kuhara (Kyushu University), Ikuko Uchiyama (Okazaki National Research Institutes), Susumu Goto, Minoru Kanehisa (Kyoto University), Naotake Ogasawara (Nara Institute of Science and Technology) and supported by the Research for the Future Program of the Japan Society for the Promotion of Science.

## FEATURES

Location/Qualifiers  
1. .347650  
/organism="Staphylococcus aureus subsp. aureus Mn50"

gene /mol\_type="genomic DNA"  
/strain="Mu50"  
/sub\_species="aureus Mu50"  
/db\_xref="taxon:158878"  
157. .1443  
CDS /locus\_tag="SAV1936"  
157. .1443  
/locus\_tag="SAV1936"  
/codon\_start=1  
/transl\_table=11  
/product="similar to aspartate transaminase protein"  
/protein\_id="BAB58098.1"  
/db\_xref="GI:14247709"  
/translation="MPLAQLNEQLQSQSNATFTMLSDLGQNMFPKGTLSQSAEAK  
STTYNATIGMATKQGMFASLSDAMDNDLTPDEIPYAPQGLEELRLWQKMLRD  
NPBLSIDMSGLPTVDSALVGLDFVNGDITLLPEHNNGLKLVFNTRNGANL  
QTYPIFDQGHYTTDSALVGLDFVNGDITLLPEHNNGLKLVFNTRNGANL  
LADQTKVIAVDDAVGLFVEDVYTSQKQVIMILNPNPNTGYTPTHKEVTIYDAIKA  
RVGPMFTGSDQTTKELEAKVGLIRSNISGSLPQTSQSAVKHVKNNKQFDEIQFN  
IQTLKERYEVKEVYADQSHHWQYDFNSGYFMAIKVHDVDPEELRKHLEKYSIG  
VIALNATDIRAFCEKEKEDIPHFVDSIAKAIDDLR"  
complement (1722. .2042)  
/locus\_tag="SAV1937"  
complement (1722. .2042)  
/locus\_tag="SAV1937"  
/note="Bacteriophage phi Sa 3mu"  
/codon\_start=1  
/transl\_table=11  
/product="truncated map-w protein"  
/protein\_id="BAB58099.1"  
/db\_xref="GI:14247710"  
/translation="MKKINKSNKVPVTKNPENIVSNFVFNASKITINDLSIK  
LKSAMANDQGITKHIDIGLAERAVYKVPKNGSSKKVDLTKYKDERVFKATDIKKVDI  
ELKP"  
complement (2027. .3457)  
/locus\_tag="SAV1938"  
complement (2027. .3457)  
/locus\_tag="SAV1938"  
/note="Bacteriophage phi Sa 3mu"  
/codon\_start=1  
/transl\_table=11  
/product="truncated map-w protein"  
/protein\_id="BAB58100.1"  
/db\_xref="GI:14247711"  
/translation="MKFKSLITTTALGVIASTGANFVNTNEASAAKPLDKSSSTLHH  
GHSNIQIPTYITVNGTSQNTLSLSSLTQNKQISYDIENKVKSVLYFNRGISDIDLRL  
SKQAEYVHFKNQTRVLDLKSIGYIADLINTSDIKASVNVDTKKQPKDKAKANQV  
PVTITVNGTSQNLISLTFNKNQISYKDLGKVKSVLESNRGITDVLRLSKQAKYT  
VNFNGTKVYIDLKSGIYTANLINSDDIKSINIVDTKKHLENKAKRYQVPYSINLN  
GTSNLSNLSFSPNFWNYKNLTSQIKSVLKHDRGISQDOLKYAKKAYTVYFKNGG  
KRILQNSKNYTNANLVHAKDVKRIETITKGTGKAKADRYPTTIAVNGTSTPILSKLK  
ISNKQLISYKLNDRKSVLKSERGISDLDLKFKAQKQKTYVYFKNGKKQVNVNLSKDIF  
TPNLFSAKDIDKIDIVQYTKSKK"  
3803. .4042  
/gene="truncated-hlb"  
/locus\_tag="SAV1939"  
3803. .4042  
/gene="truncated-hlb"  
/locus\_tag="SAV1939"  
/note="Bacteriophage phi Sa 3mu"  
/codon\_start=1  
/transl\_table=11  
/product="truncated beta-hemolysin"  
/protein\_id="BAB58101.2"  
/db\_xref="GI:46395539"  
/translation="MHLFCCKLYKRDNDGKTKNSLKKVATLALANLLVGLTNDN  
SAKESKDDTDLDLVSHNVYMLSTVLYPNWRLT"  
4414. .4593  
/locus\_tag="SAV1940"  
4414. .4593  
/locus\_tag="SAV1940"  
/note="Bacteriophage phi Sa 3mu"  
/codon\_start=1

gene /transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAB58102.1"  
/db\_xref="GI:14247713"  
/translation="MKFKTTKECKGNNIFKRSQEIINNSESEKGCWLGISMLILLFLLI  
LEGITACSSSIHFIN"  
4617. .4925  
CDS /locus\_tag="SAV1941"  
4617. .4925  
/locus\_tag="SAV1941"  
/note="Bacteriophage phi Sa 3mu"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAB58103.2"  
/db\_xref="GI:46395540"  
/translation="MKNHTNIIINILVIVNSLTHFLTNTSFNNSASDFCFIIGAIF  
FLIGFVAIYGMKRAYIWNLLLTNIFVFLHFCVLLLLKVIYGPXLFYEGCVLLFT  
"  
complement (5216. .5566)  
/locus\_tag="SAV1942"  
complement (5216. .5566)  
/locus\_tag="SAV1942"  
/note="Bacteriophage phi Sa 3mu"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAB58104.1"  
/db\_xref="GI:14247715"  
/translation="MKIRKSLAGTIAVLASPLVNLNLDKNEAQASTLSPTSNEYQNE  
KLANEKSLDELNVNELATGSLTYKRTIKISGQKAMVALKDKPFKMSKARYQLQ  
KIYNEIDEALUKSY"  
complement (6077. .6373)  
/gene="truncated-lytA"  
/locus\_tag="SAV1943"  
complement (6077. .6373)  
/gene="truncated-lytA"  
/locus\_tag="SAV1943"  
/note="Bacteriophage phi Sa 3mu"  
/codon\_start=1  
/transl\_table=11  
/product="truncated amidase"  
/protein\_id="BAB58105.2"  
/db\_xref="GI:46395541"  
/translation="MPAGYTLDRNNVPYKKTGYTYTVANVKNVNRDGYSTNSRITGV  
LPNNATIKYDQAYCINGYRWITYIANNQRCYIATGEVDKAGNRKISSGFNSAL"  
6461. .6616  
/gene="truncated-lytA"  
/locus\_tag="SAV1943.1"  
6461. .6616  
/gene="truncated-lytA"  
/locus\_tag="SAV1943.1"  
/note="Bacteriophage phi Sa 3mu"  
/codon\_start=1  
/transl\_table=11  
/product="truncated amidase"  
/protein\_id="BAD16608.1"  
/db\_xref="GI:46395542"  
/translation="MRCSDCNISMLRYMYRVKGGKINDICRISVRLGLHGLTGYIVHSY  
YIHLIIT"  
complement (7063. .7554)  
/gene="sak"  
/locus\_tag="SAV1944"  
complement (7063. .7554)  
/gene="sak"  
/locus\_tag="SAV1944"  
/note="Bacteriophage phi Sa 3mu"  
/codon\_start=1  
/transl\_table=11  
/product="staphylokinase precursor"  
/protein\_id="BAB58106.1"  
/db\_xref="GI:14247717"  
/translation="MLKRSLLFLTIVLLLFSSSITNEVSASSSFDKGYKKGGDASY

FEPTGPMVNVTVGDKGNELLSPHYVEPIKPGTTLTKIEIYVVEWALDATAYKE  
FRVLEDPASAKIEVYYDKNKKKEETKSPFITEKGFVVPDLSEHIKNPGFNLIKVI  
EKK"  
gene complement (7769. .8500)  
CDS /locus\_tag="SAV1945"  
complement (7769. .8500)

Alignment Scores:  
Pred. No.: 7.26e-81 Length: 347650  
Score: 1233.00 Matches: 238  
Percent Similarity: 84.59% Conservative: 42  
Best Local Similarity: 71.90% Mismatches: 51  
Query Match: 72.96% Indels: 1  
DB: 1 Gaps: 0

US-10-724-972A-6352 (1-335) x AP003364 (1-347650)

QY	1	GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeu 20
DB	239361	GGAGTGAACG-ATGAGAGGTCTAAAAACCTTTTAGTATATTTGGGATTAATAGTTGCCTTA 239303
QY	21	IleAlaThrAlaAlaCysGlyAsnAsnSerSerAsnSerSerLysGluSerSerLys 40
DB	239302	CTTTTAGTTGCACCTTGTGTATACGGATAATCAAGTAAAAAAGAAATCATCACTAAA 239243
QY	41	AspGlyValGluLysHisGluGluGlyThrThrLysValProLysHisProLysArg 60
DB	239242	GATACATATTTCCGTAAGAGTAAATGGTACAGTAAAGTACTTAAAGATGCAAACTG 239183
QY	61	ValValValGluLysSerPheValAspAlaLeuValAlaLeuAspValLysProVal 80
DB	239182	ATCGTTGTATTAGTACTCATTTGCAGATGTCATTAGCAGCATTAGACGTTAAACAGTT 239123
QY	81	GlyIleAlaAspAsnLysLysAsnArgIleLeuPheProLeuArgAspLysIleGly 100
DB	239122	GGTATTGCTGATGATGGTAAGAAAAACGTATCATTAACCCAGTTAGAGAAAAATTTGG 239063
QY	101	LysThrThrSerValGlyThrArgLysGlnProAsnLeuGluGluLysSerLysLeuLys 120
DB	239062	GATTATACCTCTGTAGTACACGTAACAGCCAACTTAGAGGAATTTAGTAAATTTAAA 239003
QY	121	ProAspLeuIleAlaAspAsnAsnArgHisLysGlyIleThrLysAspLeuAsnLys 140
DB	239002	CCGGATTAAATATCGCTAGTACGATAGACATAAGGATTAATAAAGAAATTTAAACAAA 238943
QY	141	IleAlaProThrIleGluLysSerPheAspGlyAspThrAsnGluAsnIleAspAla 160
DB	238942	ATTGACCAACATTTATCAATAAGAGTTTTGTAGGAGACTACAAACAAATATTAAATTCG 238883
QY	161	PheLysThrIleSerLysAlaLeuGlyLysGluGluGluGlyLysArgLeuGluGlu 180
DB	238882	TTCAAAACAATTTGCTAAAGCTTTAAATAAGAAAAAGAGCGCAAAACGCTTCTGCTGAG 238823
QY	181	HisAspLysLysIleGluGluThrLysLysGluIleThrMetAspLysAsnGlnLysVal 200
DB	238822	CATGATAAATTAATCAATAAGTATTAAGATGAATTAATTAATTTGATAGAAATCAAAAGTG 238763
QY	201	LeuProAlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerThrValGly 220
DB	238762	CTTCCAGAGTAGTTGCTAAAGCTGTTATTAGCACATCCAAACATATTATCATGTTTGA 238703
QY	221	GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAspValThrLysGlyLeu 240
DB	238702	CAATTTTAAACGAACCTAGGATTTAAAAATGCATTAAGTGACGATGTAAACAAAGGTTTA 238643
QY	241	SerLysThrLeuLysGlyProThrLysGlnMetAsnThrGluThrLeuSerGlnValAsn 260
DB	238642	AGTAAATATTGAAAGGACCTTACTTACAAATTAGACACTGAACTTTAGCTGATTTAAAT 238583
QY	261	ProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGlu 280
DB	238582	CCAGAGCGTATGATCATTTATGACAGATCATGCTAAAAAAGATTTCTGCTGAATTTCAAGA 238523

QY	281	LeuGluLysAspProValTTPlyLysLeuAsnAlaValLysAsnGlnArgValAspIle 300
DB	238522	TTTACAAGAGATGCAACATGGAATAAGTTGAATGCAGTTAAAAATATATCCGCTGGATATT 238463
QY	301	LeuAspArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLys 320
DB	238462	GTTGACCGTGATGTTTGGGCAAGATCTCGTGGCTTAATTTCTTCTGAGAAATGGCTAAA 238403
QY	321	GluLeuValGluLeuSerLysLysAspSerLys 331
DB	238402	GAACTTGTGTAATTATCAAAAAAAGAACAAAG 238370

RESULT 4  
BX571856\_23/c  
WPCOMMENT

Sequence split into 29 fragments LOCUS BX571856 Accession BX571856

Fragment Name	Begin	End
BX571856_00	1	110000
BX571856_01	100001	210000
BX571856_02	200001	310000
BX571856_03	300001	410000
BX571856_04	400001	510000
BX571856_05	500001	610000
BX571856_06	600001	710000
BX571856_07	700001	810000
BX571856_08	800001	910000
BX571856_09	900001	1010000
BX571856_10	1000001	1110000
BX571856_11	1100001	1210000
BX571856_12	1200001	1310000
BX571856_13	1300001	1410000
BX571856_14	1400001	1510000
BX571856_15	1500001	1610000
BX571856_16	1600001	1710000
BX571856_17	1700001	1810000
BX571856_18	1800001	1910000
BX571856_19	1900001	2010000
BX571856_20	2000001	2110000
BX571856_21	2100001	2210000
BX571856_22	2200001	2310000
BX571856_23	2300001	2410000
BX571856_24	2400001	2510000
BX571856_25	2500001	2610000
BX571856_26	2600001	2710000
BX571856_27	2700001	2810000
BX571856_28	2800001	2902619

Continuation (24 of 29) of BX571856 from base 2300001 (BX571856 Staphylococcus aureus su

Alignment Scores:  
Pred. No.: 2.58e-81 Length: 110000  
Score: 1231.00 Matches: 239  
Percent Similarity: 83.99% Conservative: 39  
Best Local Similarity: 72.21% Mismatches: 53  
Query Match: 72.84% Indels: 1  
DB: 1 Gaps: 0

US-10-724-972A-6352 (1-335) x BX571856\_23 (1-110000)

QY	1	GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeu 20
DB	50332	GGAGTGAACG-ATGAGAGGTCTAAAAACCTTTTAGTATATTTGGGATTAATAGTTGCCTTA 50274
QY	21	IleAlaThrAlaAlaCysGlyAsnAsnSerSerAsnSerSerLysGluSerSerLys 40
DB	50273	TTTTTAGTTGCAGCTTGTGTTAATACGATAATTTCAAGTAAAAAAGAAATCATCACTAAA 50214
QY	41	AspGlyValGluLysHisGluGluGlyThrThrLysValProLysHisProLysArg 60
DB	50213	GATACATATTTCCGTTAAAGATGAAATGGTACAGTAAAGTACCTAAAGATGCAAAACGT 50154
QY	61	ValValValLeuGluThrSerPheValAspAlaLeuValAlaLeuAspValLysProVal 80
DB	50153	ATCGTTGTATTAGTACTCATTTGCAGATGATTTGCAGCATTTAGACGTTAAACAGTTT 50094

QY	81	GlyVileAlaAaspAenLysLysAenArgIleileiyProLeuArgAaspLysleGly	100
Db	50093	GGTATTGCTGATCGTAGAAGAAACCGTATCATTTAAACCGATTAGAGAAAAATTCGG	50034
QY	101	LysTyThrSerValGlyThrArgLysGlnProAsnLeuGluGluileSerLysLeuLys	120
Db	50033	GATTATATCTCTAGTGACACGTAACACCCAACTTAGAGAAATTTAGTAATTTAAAA	49974
QY	121	ProAspLeuileileAlaAaspAenAenArgHisiLysGlyleTyriLysAepLeuAenLys	140
Db	49973	CCGGATTAAATTATCGCTGATAGCAGTAGACATAAAGGTATTAAATAAGAAATTTAAACAAA	49914
QY	141	IleAlaProThrleGluLeuLysSerPheAaspGlyAspTyraenGluAenileAaspAla	160
Db	49913	ATTGCACCAACATTATCATTTAAAGAGTTTGTAGTGACATACAGCAAAATATCGATGCT	49854
QY	161	PhelysThrileSerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGlu	180
Db	49853	TTCAAGACAATTGCGAAGCTTTAGATAAGAAAGAAAGCGGCAAAACGCTTCCTGCTGAA	49794
QY	181	HisiAepLysLysileGluGluTyriLysLysGluileThrMetAepLysAenGlnLysVal	200
Db	49793	CATGATAAATTAATCAATAAGATATAAGATGAAATTTAAATTTGATAGAAATCAAAAAAGTG	49734
QY	201	LeuProAlaValAlaLysSerGlyLeuLeuAlaHisiProSerAenSerTyriValGly	220
Db	49733	CTTCAGCAGTGTGTCGAAAGCTTTATAGCACATCCAACTATTTCGTATGTTGGA	49674
QY	221	GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAepAaspValThrLysGlyLeu	240
Db	49673	CAATTTTTTAAACCACTTGGATTTTAAATAATGCATTAAGTAGATGATGTAAACAAAGGTTTA	49614
QY	241	SerLysTyriLeuLysGlyProTyriLeuGlnMetAenThrGluThrLeuSerGlnValAsn	260
Db	49613	AGTAATACTTGAAGGACCTTACTTACAATTAGATACTGAACATTATTAGCTGACTTAAT	49554
QY	261	ProGluArgMetPheileMetThrAsnLysLysSerSerAenGlnProSerLeuLysGlu	280
Db	49553	CCTGAACGCATGATTATTATGACAGATAATGCTTAAAGAAAGATTCCTGTAATTCGAAGAG	49494
QY	281	LeuGluLysAepProValTrpLysLysLeuAenAlaValLysAenGlnArgValAspile	300
Db	49493	TTACAAGAAAGATGCAACATGGAAGAAATTTGAATGCAGTTAAATAATTCGTGTGATATT	49434
QY	301	LeuAaspArgAepLeuTrpAlaArgSerArgGlyLeuileSerSerGluGluMetAlalys	320
Db	49433	GTTGACCGTGATGTTTGGCGAAGATCTCGTGGCTTAAATTTCTCTCAAGAAATGCGTAAA	49374
QY	321	GluLeuValGluLeuSerLysLysAepSerLys 331	
Db	49373	GAACTTGTGAATTATCAAAAAAGAAACAAAG 49341	

RESULT 5	Fragment Name	Begin	End
US-10-724-972a-6352	US-10-724-972a-6352	1	110000
US-10-724-972a-6352	US-10-724-972a-6352	100001	210000
US-10-724-972a-6352	US-10-724-972a-6352	200001	310000
US-10-724-972a-6352	US-10-724-972a-6352	300001	410000
US-10-724-972a-6352	US-10-724-972a-6352	400001	510000
US-10-724-972a-6352	US-10-724-972a-6352	500001	610000
US-10-724-972a-6352	US-10-724-972a-6352	600001	710000
US-10-724-972a-6352	US-10-724-972a-6352	700001	810000
US-10-724-972a-6352	US-10-724-972a-6352	800001	910000
US-10-724-972a-6352	US-10-724-972a-6352	900001	1010000
US-10-724-972a-6352	US-10-724-972a-6352	1000001	1110000
US-10-724-972a-6352	US-10-724-972a-6352	1100001	1210000
US-10-724-972a-6352	US-10-724-972a-6352	1200001	1310000
US-10-724-972a-6352	US-10-724-972a-6352	1300001	1410000
US-10-724-972a-6352	US-10-724-972a-6352	1400001	1510000

QY	1500001	1610000	
Db	BX571857_15	BX571857_16	1710000
QY	1600001	1710000	
Db	BX571857_17	BX571857_18	1810000
QY	1800001	1910000	
Db	BX571857_18	BX571857_19	2010000
QY	1900001	2010000	
Db	BX571857_19	BX571857_20	2110000
QY	2000001	2110000	
Db	BX571857_20	BX571857_21	2210000
QY	2100001	2210000	
Db	BX571857_21	BX571857_22	2310000
QY	2200001	2310000	
Db	BX571857_22	BX571857_23	2410000
QY	2300001	2410000	
Db	BX571857_23	BX571857_24	2510000
QY	2400001	2510000	
Db	BX571857_24	BX571857_25	2610000
QY	2500001	2610000	
Db	BX571857_25	BX571857_26	2710000
QY	2600001	2710000	
Db	BX571857_26	BX571857_27	2799802
QY	2700001	2799802	

Continuation (23 of 28) of BX571857 from base 2200001 (BX571857 Staphylococcus aureus str)

Alignment Scores:	4,34e-81	Length:	110000
Pred. No.:	1228.00	Matches:	237
Score:	84.29%	Conservative:	42
Percent Similarity:	71.60%	Mismatches:	52
Best Local Similarity:	72.66%	Indels:	1
Query Match:	1	Gaps:	0
DB:			

US-10-724-972a-6352 (1-335) x BX571857\_22 (1-110000)

QY	1	GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuPheValLeu	20
Db	48090	GGAGTGGACG-ATGAGAGCTCTAAAAACCTTTTGTATATTGGGATTTAATAGTTGCCTTA	48032
QY	21	IleAlaThrAlaAlaCysGlyAenAenSerSerSerAenSerSerLysGluSerSerLys	40
Db	48031	CTTTTAGTTCAGCTTGTGTAATACGATAATTTCAAGTAAAAAGAAATCATCAACTAAA	47972
QY	41	AspGlyValGluIleLysHisGluGluGlyThrThrLysValProLysHisProLysArg	60
Db	47971	GATCTATTTCGGTAAAGATGAAATGCTACAGTAAAGTACCTAAAGATGCAAAACGT	47912
QY	61	ValValValLeuGluTyriSerPheValAspAlaLeuValAlaLeuAaspValLysProVal	80
Db	47911	ATCTGTTGTTAGTAGTACTCATTTGCAGATGCAITAGCAGCATTAGACGTTAATCCAGTT	47852
QY	81	GlyIleAlaAaspAenLysLysAenArgIleileLysProLeuArgAaspLysileGly	100
Db	47851	GGTATTGCTGATGATGTTAGAGAAAAACGATATCATTTAAACCCAGTTAGAGAAAAATTTGGG	47792
QY	101	LysTyThrSerValGlyThrArgLysGlnProAsnLeuGluGluileSerLysLeuLys	120
Db	47791	GATTATATCTCTGAGGTACAGTAAACAGCCAACTTTAGAGAAATTTAGTAAATTTAAA	47732
QY	121	ProAepLeuileileAlaAaspAenAenArgHisiLysGlyIleTyriLysAepLysAenLys	140
Db	47731	COGGATTTAATTATCGCTGATAGCAGTACATAAAGGTATTAAATAAGAAATTTAAACAAA	47672
QY	141	IleAlaProThrileGluLeuLysSerPheAaspGlyAaspTyraenGluAenileAaspAla	160
Db	47671	ATTGCACCAACATTATCATTTAAAGAGTTTGTAGTGAGACTACAAACAAACATTAATTCG	47612
QY	161	PhelysThrileSerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGlu	180
Db	47611	TTCAAAAACATTTGCTTAAGCTTTAAATAAGAAAAAGAAAGCGGAAAAACGCTCTGCTGAG	47552
QY	181	HisiAepLysLysileGluGluTyriLysLysGluileThrMetAepLysAenGlnLysVal	200
Db	47551	CATGATAAATTAATCAATAAGTATAAAGATGAAATTTAAATTTGATAGAAATCAAAAGTG	47492
QY	201	LeuProAlaValAlaAlalysSerGlyLeuLeuAlaHisiProSerAenSerTyriValGly	220
Db	47491	CTTCAGCAGTAGTGTGCTAAAGCTGTTTATTAGACATCCAAACTATTTCATATGTTGGA	47432
QY	221	GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAaspValThrLysGlyLeu	240
Db	47431	CAATTTTAAACGAACATGAGATTTTAAAGATGCAATTAAGTACGATGATCAAAAGGTTTA	47372



Qy	241	SerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsn	260
Db	47371	AGTAATATTGAAAGGACCTTACTTACAATTAGACACTGAACATTTAGCTGATTAAAT	47312
Qy	261	ProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGlu	280
Db	47311	CCAGAGCGTATGATCATTATGACTGATCAAGTAAAGAAAGATTCCTGTAATTCAGAAA	47252
Qy	281	LeuGluLysAspProValTprLysLysLeuAsnAlaValLysAsnGlnGluValAspIle	300
Db	47251	TTACAAGAAGATCAACATGGAAATAATTAATGATGATTAATAATAATTCGGCTGGATTT	47192
Qy	301	LeuAspArgAspLeuTprAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLys	320
Db	47191	GTTGACCGTATGTTGGGCAAGATCTCGTGGCTTAATTTCTTCTGAAGAATCGCTAAA	47132
Qy	321	GluLeuValGluLeuSerLysLysAspSerLys	331
Db	47131	GAACCTTGTTGAATTCAAAAAAGAACAAAAG	47099
RESULT 6	AP004829/c		
LOCUS	AP004829	304050 bp	DNA linear BCT 20-DEC-2002
DEFINITION	Staphylococcus aureus subsp. aureus MW2 DNA, complete genome,		
ACCESSION	AP004829	BA000033	
VERSION	AP004829.1	GI:21205117	
KEYWORDS			
SOURCE	Staphylococcus aureus subsp. aureus MW2		
ORGANISM	Staphylococcus aureus subsp. aureus MW2		
REFERENCE	1	Bacteria; Firmicutes; Bacillales; Staphylococcus.	
AUTHORS	Baba, T., Takeuchi, F., Kuroda, M., Yuzawa, H., Aoki, K., Oguchi, A., Nagai, Y., Iwama, N., Asano, K., Naimi, T., Kuroda, H., Cui, L., Yamamoto, K. and Hiramatsu, K.		
TITLE	Genome and virulence determinants of high virulence community-acquired MRSA		
JOURNAL	Lancet 359 (9320), 1819-1827 (2002)		
MEDLINE	22040717		
PUBMED	12044378		
REFERENCE	2	(bases 1 to 304050)	
AUTHORS	Director-General, Biotechnology Center, Aoki, K., Oguchi, A., Nagai, Y., Asano, K., Iwama, N., Baba, T., Kuroda, M., Hiramatsu, K. and Kikuchi, H.		
TITLE	Direct Submission		
JOURNAL	Submitted (06-MAR-2002) Director-General, Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center, 2chome 49-10 Nishinara, Shibuya-ku, Tokyo 151-0066, Japan (E-mail: bio@nitech.go.jp, URL: http://www.bio.nitech.go.jp/.		
FEATURES	1. 304050		
source	Location/Qualifiers		
	/organism="Staphylococcus aureus subsp. aureus MW2"		
	/mol_type="genomic DNA"		
	/strain="MW2"		
gene	/db_xref="taxon:196620"		
	complement (629..1072)		
	/gene="MW1947"		
CDS	complement (629..1072)		
	/gene="MW1947"		
	/note="ORFID:MW1947		
	hypothetical protein, similar to acetyltransferase		
	[Genomic island nu Sa 4mw]"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="BAB95812.1"		
	/db_xref="GI:21205118"		
	/translations="MSIYRFLNWSDFEKLNLCKLYDDLGYPYTNENDLKKRLAKITN		
	HDYFLLLLKKNKIKITGLSCMKMFEKNAEYMRILAPVTHSPFKKGYGKRLLAAS		
	EEFSKRLNCKAITLNSGNRRNRSLASHKLYSDNGYVNTSGGFTKQL"		
gene	complement (1178..1696)		
	/gene="MW1948"		
CDS	complement (1178..1696)		



QY 301 LeuAepArgLeuTtoAlaAerSerArgGlyLeuileSerSerGluMetAlaLys 320  
Db 170586 GTTGACCGTGTATTTGGGCAAGATCTCGTGGCTTAATTTCTTGAGAAATGGCTAAA 170527

QY 321 GluLeuValGluLeuSerLysLysAspSerLys 331  
Db 170526 GAACCTGCTGAATATCAAAAAAGACAAAAG 170494

RESULT 7  
AX622368  
LOCUS AX622368 1044 bp DNA linear PAT 20-FEB-2003  
DEFINITION Sequence 5331 from Patent WO02094868.  
ACCESSION AX622368  
VERSION AX622368.1 GI:28450503  
KEYWORDS

SOURCE  
Staphylococcus aureus  
ORGANISM Staphylococcus aureus  
Bacteria; Firmicutes; Bacillales; Staphylococcus.

REFERENCE 1  
AUTHORS Masignani, V.C., Mora, M.C. and Scarselli, M.C.  
TITLE Staphylococcus aureus proteins and nucleic acids  
JOURNAL Patent: WO 02094868-A 5331 28-NOV-2002;  
Chiron Spa (IT)

FEATURES  
Location/Qualifiers  
1..1044  
/organism="Staphylococcus aureus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:1280"

ORIGIN  
Alignment Scores:  
Pred. No.: 1,83e-82 Length: 1044  
Score: 1214.00 Matches: 237  
Percent Similarity: 84.59% Conservative: 43  
Best Local Similarity: 71.60% Mismatches: 51  
Query Match: 71.83% Indels: 2  
DB: 6 Gaps: 0

US-10-724-972A-6352 (1-335) x AX622368 (1-1044)

QY 1 GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeu 20  
Db 54 GGAGTGGAAAG-ATGAGAGGCTCTAAACCTTTT-AGTATATTGGGATTATAGTTGCCTTA 111  
QY 21 IleAlaThrAlaAlaCysGlyAsnAsnSerSerAsnSerSerLysGluSerLys 40  
Db 112 CTTTATGTTGAGCTTTGGTAAACCGGATAATTCAGCTAAAAAGATCATCAACTAAA 171  
QY 41 AspGlyValGluIleLysHisGluGlyThrThrLysValProLysHisProLysArg 60  
Db 172 GATACATTTTCGGTAAAAAGATGAAAATGTTACAGTAAAGTACCTAAAGATGCAAAACGT 231  
QY 61 ValValValLeuGlyTyrSerPheValAspAlaLeuValAlaLeuAspValLysProVal 80  
Db 232 ATCGTTGTATTAGTACTCTATTTCCAGATGCTATTAGCAGCATTAAGACGTTAAACCGTT 291  
QY 81 GlyIleAlaAspAsnLysLysAsnArgIleLysProLeuArgAspLysIleGly 100  
Db 292 GGTATTCGTGATGATGTTAAGAAAAACGATCATTAACCGTTAGAGAAAAATTCGG 351  
QY 101 LysTyrThrSerValGlyThrArgLysGlnProLeuLeuGluIleSerLysLeuLys 120  
Db 352 GATTATATCTCTGAGGTACACGTAACAGCAAACTTAGAAGAAATTAGTAATTAATA 411  
QY 121 ProAspLeuIleIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLys 140  
Db 412 CCGGATTATATTCGCTGATAGCAGTAGACATAAAGGTATTATTAAGAATTAACAAA 471  
QY 141 IleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAla 160  
Db 472 ATTCCACCAACATTTATCATTAAGAGTTTTGATGGAGACTACAAACAAATATTAAATTCG 531

QY 161 PheLysThrIleSerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGlu 180  
Db 532 TTCAAACAATTCCTAAAGCTTTAAATAAAGAAAAAGAGCGGAAAAACGCTTCCTGCTGAG 591  
QY 181 HisAspLysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysVal 200  
Db 592 CATGATAAATTAATCAATAGTATAAAGATGAAATTAATTTGATAGAAATCAAAAGTG 651  
QY 201 LeuProAlaValAlaLysSerGlyLysLeuLeuAlaHisProSerAsnSerTyrValGly 220  
Db 652 CTTCCAGCAGTAGTCTGCTAAAGCTGGTTATTAGCACATCCAAACTATTTCATATGTTGA 711  
QY 221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAspValThrLysGlyLeu 240  
Db 712 CAAATTTTAAACGAACTAGGATTTAAATAATGCATTAAAGTACCATGATTAACAAAGGTTTA 771  
QY 241 SerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsn 260  
Db 772 AGTAAATATTGAAAGGACCTTACTTACAATTAGACACTGAACATTTAGCTGATTAAAT 831  
QY 261 ProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGlu 280  
Db 832 CCAGAGCGTATGATCATTCATTATGACAGATCATGCTAAAAAGATTTCTGCTGAATTCAGAAG 891  
QY 281 LeuGluLysAspProValTyrLysLysLeuAsnAlaValLysAsnGlnArgValAspIle 300  
Db 892 TTACAAGAGATGCAACATGGAAAAAGTTGAATGCATTAATAAATAATCCGCTGGATATT 951  
QY 301 LeuAspArgAspLeuTyrAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLys 320  
Db 952 GTTGACCGTGTATTTGGGCAAGATCTCGTGGCTTAATTTCTTGAAGAAATGGCTAAA 1011  
QY 321 GluLeuValGluLeuSerLysLysAspSerLys 331  
Db 1012 GAACTTGTTGAATTTCAAAAAAGACAAAAG 1044  
RESULT 8  
BX571857.10  
WPCOMMENT

Sequence split into 28 fragments LOCUS BX571857 Accession BX571857  
Fragment Name Begin End  
BX571857\_00 1 110000  
BX571857\_01 100001 210000  
BX571857\_02 200001 310000  
BX571857\_03 300001 410000  
BX571857\_04 400001 510000  
BX571857\_05 500001 610000  
BX571857\_06 600001 710000  
BX571857\_07 700001 810000  
BX571857\_08 800001 910000  
BX571857\_09 900001 1010000  
BX571857\_10 1000001 1110000  
BX571857\_11 1100001 1210000  
BX571857\_12 1200001 1310000  
BX571857\_13 1300001 1410000  
BX571857\_14 1400001 1510000  
BX571857\_15 1500001 1610000  
BX571857\_16 1600001 1710000  
BX571857\_17 1700001 1810000  
BX571857\_18 1800001 1910000  
BX571857\_19 1900001 2010000  
BX571857\_20 200001 2110000  
BX571857\_21 2100001 2210000  
BX571857\_22 2200001 2310000  
BX571857\_23 2300001 2410000  
BX571857\_24 2400001 2510000  
BX571857\_25 2500001 2610000  
BX571857\_26 2600001 2710000  
BX571857\_27 2700001 2799802  
Continuation (11 of 28) of BX571857 from base 1000001 (BX571857 Staphylococcus aureus st

Alignment Scores:  
Pred. No.: 2.61e-43 Length: 110000

Score:	725.50	Matches:	149
Percent Similarity:	66.67%	Conservative:	63
Best Local Similarity:	46.86%	Mismatches:	95
Query Match:	42.93%	Indels:	11
DB:	1	Gaps:	4
US-10-724-972a-6352 (1-335) x BX571857_10 (1-110000)			
QY	13	ValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCysGlyAenAenSerSerSer 32	
DB	42342	GTGTGTTTATGCTAATCTTGTAGTAGCAGTAGCGGGTGTGGTCAAAAAGATACT--- 42398	
QY	33	AsnSerSerLysGluSerSerLysAspGlyValGluIleLysHisGluGluGlyThrThr 52	
DB	42399	-----GAAGAGAAAACTGAAATGACGACATAAAGATGAATAGGAAGCTGAA 42446	
QY	53	LysValProLysHisProLysAtrGValValValLeuGluTyrSerPheValAspAlaLeu 72	
DB	42447	AAATTAAGAAAAATCTCAACGTTGTTGTTATTAGAAATATAGTTTGTCTGATTATTTA 42506	
QY	73	ValAlaLeuAspValLysProValGlyLeuAlaAspAspAenLysLysAenArgIleIle 92	
DB	42507	GCAGCAATTAGATATGAAACCTGTTGGTATTGCAGATGATGCAGCAGTAAATAATATAACA 42566	
QY	93	LysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsn 112	
DB	42567	AAGTCAGTAAGAGATAAGATTGGGGCATATGAATCGGTGGATCTAGACCCCAACCGAAT 42626	
QY	113	LeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsnAsnArgHisLys 132	
DB	42627	ATGGAAGTGTAAGTAATTAATTAACCGGATTGATCATTCAGATGTTAGCAGACATAAG 42686	
QY	133	GlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGly 152	
DB	42687	AAATCAAAATCAGAAATTCAGCAAAATTCGCCGCAATATGTTAGTCAGTGGTACGGGA 42746	
QY	153	AspTyrAenGluAenIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 172	
DB	42747	GATTACAATGCAAAATTTTGAAGCATTTAAACACGTCGCTAAAGCAGTGGCAAGAGAAA 42806	
QY	173	GluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGluTyrLysLysGluIle 192	
DB	42807	GAAGCGAGAGCGTCTCGAAGAGCATGATTAATAATATTAGCGGAGATTAGAAAGAAATTT 42866	
QY	193	ThrMetAspLysAenGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAla 212	
DB	42867	GAACAGAGTACGTTAAATCTGCAATTCGCTGCTCAAGACGAGGTATGTTTATT 42926	
QY	213	HisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232	
DB	42927	AATAATGAAGATACATTTATGGGCAATCTTACTTAAATGGGTATTCAACCTGAAGTC 42986	
QY	233	SerAspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAen 252	
DB	42987	ACAAAGACAAAATACCGATGTTGGTGACCAAGGTTGTCCTTATATATATTAAAT 43046	
QY	253	ThrGluThrLeuSerGlnValAenProGluArgMetPheIleMetThrAen---LysAla 271	
DB	43047	AATGAAGACTTGCCCAATATCAATCCAAAAGTATTGATTTAGCCCACTGACCGGAAACG 43106	
QY	272	SerSerAsnGluProSerLeuLysGluLeuLysAspPro---ValTrpLysLysLeu 290	
DB	43107	GACAAAATAGAACGAAATTC-----ATTGATCCCTGCAGCTTTGGAAATCATTTA 43154	
QY	291	AsnAlaValLysAenGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArg 310	
DB	43155	AAGCTGTGAAGATTAACAAAGTTTATGACGTTGACCGAAATAGTGGTTGAAATCAAGG 43214	
QY	311	GlyLeuIleSerSerGluGluMetalLysLeuGluLeuValGluLeuSerLysLys 328	
DB	43215	GGGATTATCCCAAGTGAAAGTATGGCAGAGAATTTAGAAAAAATTTGCAGAAAAA 43268	

AP004825	290150 bp	DNA	linear	BCT 20-DEC-2002
LOCUS				
DEFINITION	Staphylococcus aureus subsp. aureus MW2 DNA, complete genome, strain,MW2, section 4/10.			
ACCESSION	AP004825	BA000033		
VERSION	AP004825.1	GI:21203989		
KEYWORDS				
SOURCE	Staphylococcus aureus subsp. aureus MW2			
ORGANISM	Staphylococcus aureus subsp. aureus MW2			
REFERENCE	Bacteria; Firmicutes; Bacillales; Staphylococcus.			
AUTHORS	1 Baba, T., Takeuchi, F., Kuroda, M., Yuzawa, H., Aoki, K., Oguchi, A., Nagai, Y., Iwana, N., Asano, K., Naimi, T., Kuroda, H., Cui, L., Yamamoto, K. and Hiramatsu, K.			
TITLE	Genome and virulence determinants of high virulence community-acquired MRSA			
JOURNAL	Lancet 359 (9320), 1819-1827 (2002)			
MEDLINE	22040717			
PUBMED	12044378			
REFERENCE	2 (bases 1 to 290150)			
AUTHORS	Director-General, Biotechnology Center, Aoki, K., Oguchi, A., Nagai, Y., Asano, K., Iwana, N., Baba, T., Kuroda, M., Hiramatsu, K. and Kikuchi, H.			
TITLE	Direct Submission			
JOURNAL	Submitted (06-MAR-2002) Director-General, Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center; 2Chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan (E-mail:bio@nitech.go.jp, URL:http://www.bio.nitech.go.jp/, Tel:81-3-3481-1933, Fax:81-3-3481-8424)			
FEATURES	Location/Qualifiers			
source	1. 290150			
organism	"Staphylococcus aureus subsp. aureus MW2"			
mol_type	"genomic DNA"			
strain	"MW2"			
db_xref	"taxon:196620"			
gene	285..1493			
CDS	/gene="MW0823"			
hypothetical protein, similar to NADH dehydrogenase"	/gene="MW0823"			
codon_start=1	/note="ORFID:MW0823"			
transl_table=11	hypothetical protein, similar to NADH dehydrogenase"			
protein_id="BAB94688.1"	/codon_start=1			
db_xref="GI:21203990"	/transl_table=11			
translation="MAODRKVLVLGAGYAGLQTVTKLOKAISTEEAEITLINKEVH	/protein_id="BAB94688.1"			
YEATLHEASAGTLNEDVLPVSEVLKDKVNEVQAEVTKIDRDAKKVETNQIYDF	/db_xref="GI:21203990"			
DILVALGFSVETFGISGMKDHAFOIENVITARELSHRHEDKFPANAKPMFSEELNHA	/translation="MAODRKVLVLGAGYAGLQTVTKLOKAISTEEAEITLINKEVH			
ILVGGAGTGVFELGELDRIPELCSKYGVDONKVKITCVEAAPKMLPMFSEELNHA	YEATLHEASAGTLNEDVLPVSEVLKDKVNEVQAEVTKIDRDAKKVETNQIYDF			
VSYLEDRGVFEKIATPIVACNEKGFVVEVDGKQQLNAGTSYWAAGVSGKLMSESE	DILVALGFSVETFGISGMKDHAFOIENVITARELSHRHEDKFPANAKPMFSEELNHA			
GVRGRIVTKQDLTNGVDNIFVIGDCSAFIPAGEERPLPTTAQIMQQGSGVAKNIK	ILVGGAGTGVFELGELDRIPELCSKYGVDONKVKITCVEAAPKMLPMFSEELNHA			
RILNGESTEEPEYVDRGTVCVSLGSHDVGVMVFDKPIAGKAAAFMKKVIDTRAVFKIG	VSYLEDRGVFEKIATPIVACNEKGFVVEVDGKQQLNAGTSYWAAGVSGKLMSESE			
IGLAFKKGKP"	GVRGRIVTKQDLTNGVDNIFVIGDCSAFIPAGEERPLPTTAQIMQQGSGVAKNIK			
1624..3099	RILNGESTEEPEYVDRGTVCVSLGSHDVGVMVFDKPIAGKAAAFMKKVIDTRAVFKIG			
/gene="ampA"	IGLAFKKGKP"			
1624..3099	1624..3099			
/gene="ampA"	/gene="ampA"			
/note="ORFID:MW0824"	1624..3099			
codon_start=1	/note="ORFID:MW0824"			
transl_table=11	codon_start=1			
product="probable cytosol aminopeptidase"	transl_table=11			
protein_id="BAB94689.1"	product="probable cytosol aminopeptidase"			
db_xref="GI:21203991"	protein_id="BAB94689.1"			
translation="MNFKLNTLSNEINTLIIGIPEHLNQLERISFNHIDITESLERL	db_xref="GI:21203991"			
KHOHLIGSKVGKIYTTAFDQVDTYRLITVGLNLTKTRSYQDMLKIWHLFYIKSEH	translation="MNFKLNTLSNEINTLIIGIPEHLNQLERISFNHIDITESLERL			
IEDTYLMDSFISKYDQSDVLMACGISERATYFDHYKSSKAPFKTNLNL1ISESL	KHOHLIGSKVGKIYTTAFDQVDTYRLITVGLNLTKTRSYQDMLKIWHLFYIKSEH			
IELDTFHGISIGISQINLARDFSNPPVLTPTQFAEDIVNHFKNTKVKVDYDVTLL	IEDTYLMDSFISKYDQSDVLMACGISERATYFDHYKSSKAPFKTNLNL1ISESL			
VSEFGLLQAAVGVGSKHKPRLVTITTINGKDKDEAFIALVKGKITVDSGYSIKTKNGM	IELDTFHGISIGISQINLARDFSNPPVLTPTQFAEDIVNHFKNTKVKVDYDVTLL			
ATMKFDMCAANVGIIEAASALQLPVMVILACAEANMINASMKMPDDVFTALSGET	VSEFGLLQAAVGVGSKHKPRLVTITTINGKDKDEAFIALVKGKITVDSGYSIKTKNGM			
VNMNTDSEGVMLADAVFYANQYQPSIVMDPATLTGAIVALGDGDDKAAAFSNSKV1	ATMKFDMCAANVGIIEAASALQLPVMVILACAEANMINASMKMPDDVFTALSGET			
LVNLIQISEGVDMFELPITATRASIKMSIDIALVNHNTGOGKALFAASFVTHPSFG	VNMNTDSEGVMLADAVFYANQYQPSIVMDPATLTGAIVALGDGDDKAAAFSNSKV1			
QTPHIFDIAGPATTNKASYNPGKPTGPMIPIVQWLKQQ"	LVNLIQISEGVDMFELPITATRASIKMSIDIALVNHNTGOGKALFAASFVTHPSFG			
3510..4826	QTPHIFDIAGPATTNKASYNPGKPTGPMIPIVQWLKQQ"			

[illegible]

JOURNAL		Patent: WO 02059148-A 301 01-AUG-2002;		
FEATURES		Cistem Biotechnologies GmbH (AT)		
Source		Location/Qualifiers		
		1. -957		
		/organism="Staphylococcus aureus"		
		/mol_type="unassigned DNA"		
		/db_xref="taxon:1280"		
ORIGIN				
Alignment Scores:		Length: 957		
Pred. No.:		1.05e-45		
Score:		724.50		
Percent Similarity:		66.67%		
Best Local Similarity:		46.86%		
Query Match:		42.87%		
DB:		6		
US-10-724-972A-6352 (1-335) x AX583924 (1-957)				
QY	13	ValileGlyLeuLeuPheValLeuileAlaThrAlaAlaCysGlyAenAenSerSerSer 32		
Db	25	GTTGTGTTTCATGCTAATTTTGTAGTACGCGGTTGTGGTCAAAAAGATACT- 81		
QY	33	AsnSerSerLysGluSerSerLysAspGlyValGluileLysHisGluGluGlyThr 52		
Db	82	-----GAAGAGAAAACTGAAATGACGACATAAAAGATGAATTAGGACTCAA 129		
QY	53	LysValProLysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeu 72		
Db	130	AAATTAAGAAAAATCCTAAACCGTGTGTGTATTAGATATAGATTTTGTCTGATTATTA 189		
QY	73	ValAlaLeuAspValLysProValGlyileAlaAspAspAenLysLysAenArgIleile 92		
Db	190	GCAGCATTAGATATGAACCTGTTGGTATTGCAGATGATGCGACACCTAAATAATATAACA 249		
QY	93	LysProLeuArgAspLysLysLysGlyLysTyrThrSerValGlyThrArgLysGlnProAen 112		
Db	250	AAATCAAAATCAGAAATGAGCAAAATGCTCCGCAATCATGTTAGTTCGGGTACGGGA 429		
QY	133	GlyileTyrLysAspLeuAenLysLysLysLysLysLysLysLysLysLysLysLys 152		
Db	370	AAATCAAAATCAGAAATGAGCAAAATGCTCCGCAATCATGTTAGTTCGGGTACGGGA 429		
QY	153	AspTyrAenGluAenLysLysLysLysLysLysLysLysLysLysLysLysLysLys 172		
Db	430	GATTATATGCAAAATATTGAAGCATTTTAAACAGCTCGCTAAAGCAGTAGGCAAGAGAA 489		
QY	173	GluGlyLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 192		
Db	490	GAAGCGGAGAGCGCTCGTGAAGCATGTTAAATATTAGCGGAGATTAGAAAGAAATTT 549		
QY	193	ThrMetAspLysAenGlnLysValLeuProAlaValAlaAlaLysSerGlyLeuLeuAla 212		
Db	550	GAACAGAGTAGCTTAAATCTGCATTTGCTATTCCTCAAGACGAGGTATGTTATT 609		
QY	213	HisProSerAenSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232		
Db	610	AATAATGAAGATACATATTTATGGGCAATTTCTTAATTTAAATGGGTATTCACCTGAGTC 669		
QY	233	SerAspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAen 252		
Db	670	ACAAAGAGCAAAACTACGCATGTTGGTGAACCGCAAGGTGGTCTTATATATATATTAAT 729		
QY	253	ThrGluThrLysSerGlnValAenProGluArgMetPheIleMetThrAen---LysAla 271		
Db	730	AATGAAGAACTTGCCCAATATCAATCCAAAAAGTATGATTTAGCCACTACGCGAAACG 789		
QY	272	SerSerAenGluProSerLeuLysGluLeuLysAspPro---ValTrpLysLysLeu 290		

QY	53	LysValProLysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeu 72	
Db	108124	AAATTAAGAAAAATCCTAAACCGTGTGTGTATTAGATATAGATTTTGTCTGATTATTA 108183	
QY	73	ValAlaLeuAspValLysProValGlyileAlaAspAspAenLysLysAenArgIleile 92	
Db	108184	GCAGCATTAGATATGAACCTGTTGGTATTGCAGATGATGCGACAGTAAATAATATAACA 108243	
QY	93	LysProLeuArgAspLysLysLysLysLysLysLysLysLysLysLysLysLysLys 112	
Db	108244	AAATCAAAATCAGAAATGAGCAAAATGCTCCGCAATCATGTTAGTTCAGTGTACGGGA 108303	
QY	113	GluGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 132	
Db	108304	ATGGAAGTGATGAATTAATTAAGCATTTTAAACAGCTCGCTAAAGCAGTTGCAACAGATAAG 108363	
QY	133	GlyileTyrLysAspLeuAenLysLysLysLysLysLysLysLysLysLysLysLys 152	
Db	108364	AAATCAAAATCAGAAATGAGCAAAATGCTCCGCAATCATGTTAGTTCAGTGTACGGGA 108423	
QY	153	AspTyrAenGluAenLysLysLysLysLysLysLysLysLysLysLysLysLysLys 172	
Db	108424	GATTACAAATGCAATATTGAAGCATTTTAAACAGCTCGCTAAAGCAGTTGCAACAGATAAG 108483	
QY	173	GluGlyLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 192	
Db	108484	GAAGCGGAGAGCGCTCGTGAAGCATGTTAAACAGCTCGCTAAAGCAGTTGCAACAGATAAG 108543	
QY	193	ThrMetAspLysAenGlnLysValLeuProAlaValAlaAlaLysSerGlyLeuLeuAla 212	
Db	108544	GAACAGAGTAGCTTAAATCTGCATTTGCTATTCCTCAAGACGAGGTATGTTATT 108603	
QY	213	HisProSerAenSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232	
Db	108604	AATGAAGAACTTGCCCAATATCAATCCAAAAAGTATGATTTAGCCACTACGCGGAAACG 108663	
QY	233	SerAspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAen 252	
Db	108664	ACAAAGAGCAAAACTACGCATGTTGGTGAAGCATGTTGCTTATATATATATATTAAT 108723	
QY	253	ThrGluThrLysSerGlnValAenProGluArgMetPheIleMetThrAen---LysAla 271	
Db	108724	AATGAAGAACTTGCCCAATATCAATCCAAAAAGTATGATTTAGCCACTACGCGGAAACG 108783	
QY	272	SerSerAenGluProSerLeuLysGluLeuLysAspPro---ValTrpLysLysLeu 290	
Db	108784	GACAAAAATAGAACCAATTC-----ATTGATCCTGCAGTTTGGAAATCATTA 108831	
QY	291	AenAlaValLysAenGlnArgValAspLysLysLysLysLysLysLysLysLysLys 310	
Db	108832	AAAGCTGTGAAGATAAACAAAGTTTATGACGTTGACCGCAATTAAGTGTGAAATCAAGG 108891	
QY	311	GlyLeuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 328	
Db	108892	GGGATTATCGCAAGTGAAAGTATGCGAAGAGATTAGAAAAAATTTGCAAAAAA 108945	

RESULT 10  
AX583924  
LOCUS  
DEFINITION  
Sequence 301 from Patent WO02059148.  
ACCESSION  
AX583924  
VERSION  
AX583924.1 GI:27655594  
KEYWORDS  
Staphylococcus aureus  
SOURCE  
Staphylococcus aureus  
Bacteria; Firmicutes; Bacillales; Staphylococcus.  
REFERENCE  
1  
Meinke, A., Nagy, E., von Ahsen, U., Klade, C., Henics, T., Zauner, W.,  
Minh, D. B., Vytvytska, O., Eitz, H., Dryla, A., Weichhart, T., Hafner, M.,  
Tempelmeier, B., Fraser, C. M. and Gill, S.  
A method for identification, isolation and production of antigens  
to a specific pathogen

```
Db 790 GACAAAATAGAACCAATTC-----ATTGATCTCGCAGTTTGGAAATCATTA 837
QY 291 AsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrrAlaArgSerArg 310
Db 838 AAGCTGTGAAGAGATAACAAGTTTATGACGTTTGACCGAAATAGTGGTTGAAATCAAGG 897
QY 311 GlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328
Db 898 GGGATTATCCAGTGAAGATATGGCAGAGATTATAGAAAAAATTCGAGAAAAA 951

RESULT 11
AX621226
LOCUS 724.50 975 bp DNA linear PAT 20-FEB-2003
DEFINITION Sequence 4189 from Patent WO02094868.
ACCESSION AX621226
VERSION AX621226.1 GI:28449932
KEYWORDS
SOURCE Staphylococcus aureus
ORGANISM Staphylococcus aureus
REFERENCE 1
AUTHORS Masignani, V.C., Mora, M.C. and Scarselli, M.C.
TITLE Staphylococcus aureus proteins and nucleic acids
JOURNAL Patent: WO 02094868-A 4189 28-NOV-2002;
Chiron Spa (IT)
FEATURES
Location/Qualifiers
source 1..975
/organism="Staphylococcus aureus"
/mol_type="unassigned DNA"
/db_xref="taxon:1280"

ORIGIN
Alignment Scores:
Pred. No.: 1,07e-45 Length: 975
Score: 724.50 Matches: 149
Percent Similarity: 66.67% Conservative: 63
Best Local Similarity: 46.86% Mismatches: 95
Query Match: 42.87% Indels: 11
DB: 6 Gaps: 4

US-10-724-972A-6352 (1-335) x AX621226 (1-975)

QY 13 ValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCysGlyAsnSerSerSer 32
Db 43 GTTGTTTCATGCTAATTTAGTTAGCAGTAGCGGGTTGGTCAAAAGATACT--- 99
QY 33 AsnSerSerLysGlySerSerLysAspGlyValGluIleLysHisGluGluGlyThr 52
Db 100 -----GAGAGAGAACTGAATGACGACACATAAAGATGAATTAGGAATCTGAA 147
QY 53 LysValProLysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeu 72
Db 148 AAAATTAAAGAAAAATCCTAAACGCTGTTGTTATTAGAAATATAGTTTTCGCTGATTATTA 207
QY 73 ValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLysLysAsnArgIlele 92
Db 208 GCAGATTAGATATGAACCTGTTGGTATTGCAGATGTCGACACTTAAATAATATACA 267
QY 93 LysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGluProAsn 112
Db 268 AAGTCAGTAAGACATAAGATTGGGCGCATATGAATCGTTGGTATCAGACCCGACCGAAT 327
QY 113 LeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsnAsnArgHisLys 132
Db 328 ATGGAAGTGTAAGTAAGTAATTAATTAACCGGATTGATCATTCAGATGTTAGCAGACATAAG 387
QY 133 GlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGly 152
Db 388 AAATCAATCAGAAATGAGCAAAATTCCTCCGCAATCATGTAGTTAGCGGTACGGGA 447
QY 153 AspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 172
Db 448 GATTATAATGCAATATTGAAGCATTTAAACACAGTCGCTAAGCAGTAGGCAAGAGAAA 507
```

```
QY 173 GluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGluTyrLysLysGluIle 192
Db 508 GAAGCGCAGAGAGCGTCTGGAAGACATGATAAATATTAGCGGAGATTAGAAAGAAATTT 567
QY 193 ThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAla 212
Db 568 GAACAGAGTAGCTTAAATCTGCATTTGCATTTCGTTATCTCAAGAGCAGGTATGTTTATT 627
QY 213 HisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232
Db 628 AATAATGAGATACATATTATGGGACAAATCTTAATTAATGGGTATTCAACCTCAAGTC 687
QY 233 SerAspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsn 252
Db 688 ACAAAAGACAAACTACGCATGTTGCTGAACCGAGGTGGTCTTATATATATTAAAT 747
QY 253 ThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn---LysAla 271
Db 748 AATGAAGAACTTGCCTCAATATCAATCCAAAGTTATGATTTTAGCCACTGACGGAACG 807
QY 272 SerSerAsnGluProSerLeuLysGluLeuGluLysAspPro---ValTrpLysLysLeu 290
Db 808 GACAAAATAGAACCAATTC-----ATTGATCTCGCAGTTTGGAAATCATTA 855
QY 291 AsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrrAlaArgSerArg 310
Db 856 AAAGCTGTGAAGAGATAACAAGTTTATGACGTTTGACCGAAATAGTGGTTGAAATCAAGG 915
QY 311 GlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328
Db 916 GGGATTATCGCAAGTGAAGATATGCGAGAAGATTTAGAAAAAATTCGACAAAAA 969
```

```
RESULT 12
AP003132 298050 bp DNA linear BCT 11-JAN-2003
LOCUS Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
DEFINITION genome, section 4/10.
ACCESSION AP003132 BA000018
VERSION AP003132.2 GI:14349174
KEYWORDS
```

```
SOURCE Staphylococcus aureus subsp. aureus N315
ORGANISM Staphylococcus aureus subsp. aureus N315
REFERENCE 1
AUTHORS Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I., Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Lian, J., Ito, T., Kanamori, M., Matsumaru, H., Maruyama, A., Murakami, H., Hosoyama, A., Mizutani, U., Y., Takahashi, N.K., Sawano, T., Inoue, R., Kaito, C., Sekimizu, K., Hirakawa, H., Kuhara, S., Goto, S., Yabuzaki, J., Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C., Shiba, T., Hattori, M., Ogasawara, N., Hayashi, H. and Hiramatsu, K.
TITLE Whole genome sequencing of methicillin-resistant Staphylococcus aureus
```

```
JOURNAL Lancet 357 (9264), 1225-1240 (2001)
MEDLINE 21311952
PUBMED 11418146
REFERENCE 2 (bases 1 to 298050)
AUTHORS Director-General, Biotechnology Center, Aoki, K., Oguchi, A., Hosoyama, A., Nagai, Y., Kuroda, M., Hiramatsu, K. and Kikuchi, H.
TITLE Direct Submission
JOURNAL Submitted (30-JAN-2001) Director-General, Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center, 2Chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan (E-mail: bioelite.go.jp, URL: http://www.bio.nite.go.jp/, Tel: 81-3-3481-1933, Fax: 81-3-3481-8424)
COMMENT On Jun 12, 2001 this sequence version replaced gi:13700734.
```

```
FEATURES
Location/Qualifiers
source 1..298050
/organism="Staphylococcus aureus subsp. aureus N315"
/mol_type="Genomic DNA"
/strain="N315"
/sub_species="aureus N315"
```

gene  
CDS  
/db\_xref="taxon:158879"  
280..432  
/gene="SA0792"  
280..432  
/gene="SA0792"  
/note="ORFID:SA0792"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAB42031.1"  
/db\_xref="GI:13700735"  
/translation="MKSQKPPNKKVVEAFKPYLLTLLYLAIFITLYLYGGSDTHNN  
FYNEP"  
448..1905  
/gene="dltA"  
448..1905  
/gene="dltA"  
/note="ORFID:SA0793"  
/codon\_start=1  
/transl\_table=11  
/product="D-alanine-D-alanyl carrier protein ligase"  
/protein\_id="BAB42032.1"  
/db\_xref="GI:13700736"  
/translation="MTDINKLQAFADANPQSIARHTTDELTYQQLMDESSKLAHLR  
QSKKPMILGHMSPYMLVGMIGAIKAGCGVVPVDTSTIPEDRIKMIINKVQPEFVNT  
TDSFSELEGEVFTIEDIKTSQDPVIFDSQIKONDVTYTLFTSGTGEPKGVQLEYAS  
LVQETWMLNLSGNQOQMLNQAPFSDLSVMAIYPCLASGGTLNLVDKMNKPKL  
NEMLTATPINIWSTPSFMEMCLLLPTLINEEOYGSNEFFFCGEILLPHRAAKALVR  
PNSATINTVGTATVAVT"SIQITQILDOYPTLPVGEVRLGARLSTDDGELVIEG  
QSVSLGKNDOKTAEVFNFDGIRTVHTGDKAKFENGQWFIQGRIDFQIKLNGYRME  
LBEETQLRQSEFKEALVVPVKNDKVIHLIGAIVPTTEVTDNAENTKNIKNDLKR  
LPEYMIPRKFEWMEQLPLTNSGKIDRKKIAEVING"  
1902..3116  
/gene="dltB"  
1902..3116  
/gene="dltB"  
/note="ORFID:SA0794"  
/codon\_start=1  
/transl\_table=11  
/product="DltB membrane protein"  
/protein\_id="BAB42033.1"  
/db\_xref="GI:13700737"  
/translation="WIPXGDTFFELIALLPVILFGLGKRSYVINGVVTAFMVLV  
IPSDKHNLFQOKYLSVQLISFIYVWQVLLIMPHYHSPKNNSESKFVTVMVLSIL  
PLALVQLVQWLGQDHFHESKLIEFVGLGISYVTFKSVQLIMEIRDGSIEKIIV  
WKLQFISFFPTISGGPDRYKRFVKDDKVPVTGNEYRELKAIHIMMLGLFLYKIIV  
AVFTNYAIMQLDLQHLGNLWLYMAYSLYLFDFDAGYSLFAIAFSLYFKITPPN  
PDKPKAKNIQDPNRMWMTLSFWRPDCIYMRSLFYMGRKLLKKSQFAMSNVFLNF  
FIMGHWHGIEVYIYVGLYHAALFIGYGYERWRKKHPRPWQNGFTTALSIVITFHFV  
TFGLIFSGKLI"  
3134..3370  
/gene="dltC"  
3134..3370  
/gene="dltC"  
/note="ORFID:SA0795"  
/codon\_start=1  
/transl\_table=11  
/product="D-alanyl carrier protein"  
/protein\_id="BAB42034.1"  
/db\_xref="GI:13700738"  
/translation="MEFREQVNLAEVAENDIVKENPDVEIFEEGIIDSFQTVGLLL  
EIQNKLDIEVSIWDFDRDEWATPNKIVEALELR"  
3367..4542  
/gene="dltD"  
3367..4542  
/gene="dltD"  
/note="(glycerophosphate chain)  
ORFID:SA0796"  
/codon\_start=1  
/transl\_table=11  
/product="poly D-alanine transfer protein"  
/protein\_id="BAB42035.1"  
/db\_xref="GI:13700739"

gene  
CDS  
/translation="MKLKPFLILISGAVFIVFLLLPASWFTGLVNEKTVEDNRTSLT  
DQVLKGTLIQDKLYESNKYPIYSGSELGKDDPFNPAIALNKHANKKAFLLCAGGST  
DLINAVELASQYDLKGGKLTPIISQWFTNHLNQNFDARMSQTIQINOMFOCKNMS  
TELKRYAORLLQFPHVHNKEYLKSAYKNPKETKDSYISGFENQILKIETAKSLFAM  
DKSPLHVKPAKPDASWDMKQKAVEIGKADTTSNKFGIRDOYMLIIQESKFKVRD  
YEFVNSPEFQDLLELLYKTRAAQDVVVSIPSGVWYDHIGIDKERRQAVYKKIHS  
TVVDGGKIYDMTDKDIYKIVISDAVHIGKGVWYDEQIAKHKMGEPQPEVDKPKN"  
complement(4805..5047)  
/gene="SA0797"  
complement(4805..5047)  
/gene="SA0797"  
/note="ORFID:SA0797  
hypothetical protein, similar to nitrogen fixation protein  
NifU"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAB42036.1"  
/db\_xref="GI:13700740"  
/translation="MPTEDTMDQVAEVIERLPELLRDGGDCSLIDVEDGIVKLQL  
HGACGTCPSSTITLKAGIERALHEEVPVIEVEQVF"  
5163..5471  
/gene="SA0798"  
5163..5471  
/gene="SA0798"  
/note="ORFID:SA0798"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="BAB42037.1"  
/db\_xref="GI:13700741"  
/translation="MVVYGDVTCASCVNAPTSKDIYDMLQPLLKRYKYPNISPKYTVI  
DITKDNDNLHDHLOFTIERIEQDELFPYLTIMNDEYVADGYIQTKTRFIDQKLVE  
"  
complement(5530..6594)  
/gene="SA0799"  
complement(5530..6594)  
/gene="SA0799"  
/note="ORFID:SA0799  
hypothetical protein, similar to NADH dehydrogenase"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAB42038.1"  
/db\_xref="GI:13700742"  
/translation="MKNLVLLGGYGNNRIMSRILTSLPDYTVTLVDRMPFHGLKP  
BFYALAGTQDKDVRMKFPNQVNTVYGEINDIDLDQAIVSVGNSKIDYDELIGL  
GCEQHNVPQAEYTHSIOTLSKARDTFHSISELPEKAGVIGVAGLSGIELASBLR  
ESRDLELYLDRGPRIILRNEPEKLSYKAKVAKNNVTVVPSNININKEVPEKICND  
EPKIDIDLVTWTAGIQPVEVVRNLPIDINSNGRVLVNOHVQPTVYRNVVVGDCADLPH  
APSAQLAEVQGDIDVLLKQWLNEPLDKMPELKVQGIUGSLGDKQGFAYIMDRVT  
GRLASILKSGVLWLYKYHNG"  
6911..7147  
/gene="SA0800"  
6911..7147  
/gene="SA0800"  
/note="ORFID:SA0800"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="BAB42039.1"  
/db\_xref="GI:13700743"  
/translation="WNPVIFCLSNWAKGQDVFNQLENDPDVDVLEYGLTHCGICS  
AGLYALVNGDIVEGDSPELLQNIYAHIKETWIF"  
7160..7519  
/gene="SA0801"  
7160..7519  
/gene="SA0801"  
/note="ORFID:SA0801"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="BAB42040.1"  
/db\_xref="GI:13700744"  
/translation="WPTVILTEAAAYEVKDKMLKANEMPDGYLKIKNVGGCGTGLTYGM



GABEAPGENDEVLEYFGLKVLVDKEDAPVLNGTTIDPFQSLMGGGFQIDNPNAISCG  
CGSSFRKATKAGNPENC"  
7973. 9181  
/gene="SA0802"  
CDS 7973. 9181  
/gene="SA0802"  
/note="ORFID:SA0802  
hypotheical protein, similar to NADH dehydrogenase"  
/codon start=1  
/transl table=11  
/protein\_id="BAB42041.1"  
/db\_xref="GI:13700745"  
/translation="MAQDRKKVLVAGYAGLQTVTKLQKAIETBAEITLKNKNEYH  
YEATWLHEASAGTLANTEDLVYVESVLKKDKNFVQAEVTKIDRDARKVETNGQIDYF  
DILVVALGFVSETFEGMKDHFQIENVITARELSRHIEDKFANYAASKEDNDLS  
ILVGGAGFTGVFELBLDRIPELCSKYGVQDNKVIITVEAPKPLMPFSEELVNHHA

Alignment Scores:  
Pred. No.: 1.03e-42 Length: 298050  
Score: 724.50 Matches: 149  
Percent Similarity: 66.67% Conservative: 63  
Best Local Similarity: 46.86% Mismatches: 95  
Query Match: 42.87% Indels: 11  
DB: 1 Gaps: 4

US-10-724-972a-6352 (1-335) x AP003132 (1-298050)

QY 13 ValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCysGlyAsnAsnSerSerSer 32  
DB 113875 GTTGTTTATGCTAATCTTAGTTAGCAGTAGCGGTTGGTCAAAAAGATACT--- 113931

QY 33 AsnSerSerLysGlySerLysAspGlyValGluIleLysHisGluGluGlyThrThr 52  
DB 113932 -----GAAGAGAAACTGAAATGACGACATAAAGATGAATAGAACTGAA 113979

QY 53 LysValProLysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeu 72  
DB 113980 AAAATTAAGAAAAATCCTAAACGTTGTTGTTATTAGATATAGTTTCTGCTATTATTA 114039

QY 73 ValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLysLysAsnArgIleIle 92  
DB 114040 GCAGCATTAGATATGAACCTGTTGGTATTGCGAGATGATGCGCAGCACTAAAAATATAACA 114099

QY 93 LysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsn 112  
DB 114100 AAGTCAGTAAGAGATAAGATTGGGCGCATATGAATCGTTGGTATCTAGACCGCAACCGAAT 114159

QY 113 LeuGluGluIleSerLysLeuLysProAspLeuIleIleAlaAspAsnAsnArgHisLys 132  
DB 114160 ATGGAAGTCATAAGTAATAATTAACCGGATTTGATCATTTGCAGATGTTAGCAGACATAAG 114219

QY 133 GlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGly 152  
DB 114220 AAAATCAAATCAGAAATTGAGCAAAATTCGTCGACAATCATGTTAGTTAGCGGTACGGGA 114279

QY 153 AspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 172  
DB 114280 GATTATATGCATATATTGAAGCATTTAAACAGTCGCTAAAGCAGTAGGCAAGAGAGAAA 114339

QY 173 GluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGluTyrLysLysGluIle 192  
DB 114340 GAAGCGGAGAGCGCTCTGAAAGCATGATATAATATTAGCGGAGATTAGAAAGAAATTT 114399

QY 193 ThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAla 212  
DB 114400 GAACAGAGTACGTTAAATAATCTGCATTTTCGATCTCAAGAGCAGGTATGTTTATT 114459

QY 213 HisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232  
DB 114460 AATAATGAGATACATTTATGGGCAATCTTAATTAAGATTTGGGTATTTCACACCTGGAAGTC 114519

QY 233 SerAspAspValThrLysGlyLeuSerLysTyrLysLysGlyProTyrLeuGlnMetAsn 252  
DB 114470 AATAATGAGATACATTTATGGGCAATCTTAATTAAGATTTGGGTATTTCACACCTGGAAGTC 114519

Db 114520 ACAAAAGACAAAACACTAGCATGTTGGTGAACGCAAGGTGGTCTTATATATATATTTAAAT 114579

QY 253 ThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn---LysAla 271  
Db 114580 AATGAAGAAGCTTCCCAATATCAATCCAAAGTTATGATTTTAGCCACTGACGGAACACG 114639

QY 272 SerSerAsnGlnProSerLeuLysGluLeuGluLysAspPro---ValTrpLysLysLeu 290  
Db 114640 GACAAAATAAGAACGAATTC-----ATTGATCTGCAGTTTGGAAATCATTA 114687

QY 291 AsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArg 310  
Db 114688 AAAGCTGTGAAGATACAAAGTTTATGAGCTTGACCGAAATAAGTGGTTGGAATCAAGG 114747

QY 311 GlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328  
Db 114748 GGGATTATCCCAAGTCAAGATGATGCGAAGATTAGAAAAAATTCAGAAAAA 114801

RESULT 13  
AP003361 347785 bp DNA linear BCT 17-APR-2004  
LOCUS Staphylococcus aureus subsp. aureus Mu50 complete genome,  
DEFINITION section 4/9.  
ACCESSION AP003361 BA000017  
VERSION AP003361.3 GI:46395333  
KEYWORDS  
SOURCE Staphylococcus aureus subsp. aureus Mu50  
ORGANISM Staphylococcus aureus subsp. aureus Mu50  
REFERENCE 1  
Bacteria; Firmicutes; Bacillales; Staphylococcus.  
AUTHORS Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,  
Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,  
Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,  
Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,  
Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,  
Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,  
Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.  
Whole genome sequencing of methicillin-resistant Staphylococcus  
aureus  
JOURNAL Lancet 357 (9264), 1225-1240 (2001)  
MEDLINE 21311952  
PUBMED 11418146  
REFERENCE 2  
Ohta,T., Hirakawa,H., Morikawa,K., Maruyama,A., Inose,Y.,  
Yamashita,A., Oshima,K., Kuroda,M., Hattori,M., Hiramatsu,K.,  
Kuhara,S. and Hayashi,H.  
Nucleotide Substitutions in Staphylococcus aureus Strains, Mu50,  
Mu3 and N315  
DNA RESEARCH 11, 51-56 (2004)  
3 (Bases 1 to 347785)  
Hattori,M., Yamashita,A., Oshima,K. and Shiba,T.  
Direct Submission  
Submitted (28-FEB-2001) Masahira Hattori, Kitasato Institute for  
Life Sciences, Kitasato University; Kitasato 1-15-1, Sagamihara,  
Kanagawa 228-8555, Japan  
(E-mail:hattori@genome.its.kitasato-u.ac.jp, Tel:81-42-778-8194,  
Fax:81-42-778-8193)  
On Apr 15, 2004 this sequence version replaced gi:14246761.  
This work was done in collaboration with Toshiko Ohta, Mutsumi  
Kanamori, Hideo Hayashi (University of Tsukuba), Hideki Hirakawa,  
Satoru Kuhara (Kyushu University), Ikuo Uchiyama (Okazaki National  
Research Institutes), Susumu Goto, Minoru Kanehisa (Kyoto  
University), Naotake Ogasawara (Nara Institute of Science and  
Technology), and supported by the Research for the Future Program of  
the Japan Society for the Promotion of Science.  
FEATURES  
Location/Qualifiers  
1. 347785  
/organism="Staphylococcus aureus subsp. aureus Mu50"  
/mol\_type="genomic DNA"  
/strain="Mu50"  
/sub\_species="aureus Mu50"  
/db\_xref="taxon:158878"  
11. .997  
gene

/locus\_tag="SAV0992"
11..997
/locus\_tag="SAV0992"
/codon\_start=1
/transl\_table=11
/product="oligopeptide ABC transporter ATP-binding protein homologue"
/protein\_id="BAB57154.1"
/db\_xref="GI:14246762"
/translations="MNNVLEVDKLETSKINNEWLATVENISFELSKGEVLGIVGES
CGCKSLTSKIIKLLPEKISKLSSGEVIFDGKRIIDTINEKQLDIDRGNDAIMIFQBP
TALNPVFTVKNOLVESIKSHKSKKEANNLAKDLKKVGIARQDEILNSYFHLGGG
MRQRVMAISCSPLLIADDEPTTALDVTIOAIDLLKELOKETOMAIMTHTDLS
VVAEFCDDKLVIVAGQIVFEGGKEIHLNPKHYTKLSTIPKLEGGOKLELTIEGI
VPSIQAFHNKCRFANRCKKLDICNNQSPMHVCEVDIVRCHLYKNEYKEI"
1000..1980
/gene="appf"
/locus\_tag="SAV0993"
1000..1980
/gene="appf"
/locus\_tag="SAV0993"
/codon\_start=1
/transl\_table=11
/product="oligopeptide transport system ATP-binding protein Appf homologue"
/protein\_id="BAB57155.1"
/db\_xref="GI:14246763"
/translations="MENILEVNOIKKYYIKTGLLQKTVKVAVDVSVFSIKKGQTFG
LVGSGCGKTLGVIRLEDATSGSIIIVNGEDITRLQGGKKRSQOYQIMIFQDPYA
SLNPMQVGIISPELILNYKKLPKEIKKEVLYLLKCVGLSEDAYYKYAEHFGSGQ
RVGARALALPLSIVADEPVSALDVSVQSVQNLKLOEQPKLSYLFTHADLSVVK
HISDVIQVYMLGHVIEASDKIYEENPKHYTKLSTIPKLEGGOKLELTIEGI
PSNPQCGPTRLTRCFIADKCKENIPQLKDIGNEHQVACFYVDKVGDLNA"
1973..2935
/gene="oppb"
/locus\_tag="SAV0994"
1973..2935
/gene="oppb"
/locus\_tag="SAV0994"
/codon\_start=1
/transl\_table=11
/product="probable oligopeptide transport system permease protein"
/protein\_id="BAB57156.1"
/db\_xref="GI:14246764"
/translations="MLKILKRLGLMIPLLIISIVFSLAIQDPFSLQNGKIK
QEAINAQKGLNDISPHOYTRWNVHVHGDGSEIKYKRPVIDVIERPNTILLG
AMSLIITVISPALGITSGRYSYSLTDTVOIFENYLMIAIPSIAGVFAIFISFELQ
WPPQGSVDINIKEGTFPYNSKIYHTPLPALTGLLSTAGVIOYLRNDIENSCKDY
VLTARSKGLSMNKIYNKHLRNLSPITITFLGADIVSILGAVITETIFSNGICKLF
LESVTGQDYPPLMMALTLPFSFLGLGNLISDITYGFIIDPRNSN"
2947..3828
/locus\_tag="SAV0995"
2947..3828
/locus\_tag="SAV0995"
/codon\_start=1
/transl\_table=11
/product="similar to oligopeptide ABC transporter permease"
/protein\_id="BAB57157.1"
/db\_xref="GI:14246765"
/translations="MWNKSKSPFKIALSRFIHNKIAMLSIIFLLIITVSIAPLIAP
FPVQDNLNIGKEMTAQNLGDSGRDNFRLLYAGRISLSIGTITGMLLIGIT
VGVISGFGGIVDTLLMIRITEFVMLPFLIFAIVLNAAIGDKIKNPYGSAILVLVII
VLWSGIIARLGVKLOKENEYFLAASKISGTPYTKIILKLLPLLSVIVQATLLF
AGMIVVESGLSFLGFGISKAI19WGNMLSDAQEGDVISGPKMIMPEAITMITLISL
NPFVGEGLKDAFPNPRGR"
complement(3870..4859)
/gene="trps"
/locus\_tag="SAV0996"
complement(3870..4859)
/gene="trps"
/locus\_tag="SAV0996"

/codon\_start=1
/transl\_table=11
/product="tryptophanyl-tRNA synthetase"
/protein\_id="BAB57158.1"
/db\_xref="GI:14246766"
/translations="METIPSGIQPSGIPTIGNVIGALKOFVDVNDYDCYFICVDQHA
ITMPQRLKRLQRLAAIYLASGIDPDKATILFQSEVPAHVQAGMMITTIASVCEL
ERMTQYKQAKAVEGIPAGLLTYPLMAADIVLNTNINVPVGGDDQKHIELTRNLVD
RNSRYNDVLVKEIPRMPGGRVMSLQDPTPRKMSKSDNNAKNFISLLDEPNVAARKI
KAFVTSDDGIIKPDNDKPGITNLIISIYAGLTDMPIKOIEAKYEGEGYKGFKGLDAEI
VKAFLVFOEKYESFVNSDKLDDIILDQGRDKAHKVSFKTVKKMEKAMGLGRKR"
5153..5548
/locus\_tag="SAV0997"
5153..5548
/locus\_tag="SAV0997"
/codon\_start=1
/transl\_table=11
/product="arsenate reductase family protein"
/protein\_id="BAB57159.1"
/db\_xref="GI:14246767"
/translations="MVTLFSPSCTSCRKAKAWLQBDHPIPYTERNIFSEHLTIDEIKQ
ILKMTEDGTDEIISTRSKYQKLNVDISLPLQDLYSIIQDNFGLLRPIIDNKRILQ
VGNDEIRPLPRKYRTTQLQEAQRMD"
5919..6638
/locus\_tag="SAV0998"
5919..6638
/locus\_tag="SAV0998"
/codon\_start=1
/transl\_table=11
/product="similar to negative regulator of genetic competence Meca"
/protein\_id="BAB57160.1"
/db\_xref="GI:14246768"
/translations="MRIERVDVTKVLFTIYSIDIEARGFSGREDLWNRKGESEFWSM
MDINEBEDFVVEGPIWOVHAPEKGEVETISKSNEDMMNMSDDDATDQFDEQVQL
LAQTEGEDQLEELFEORTKEKAQSKRKQSSARKNTRTIIIVKNFLEVDVINYAHS
NPITTEFDLLYMVDTGYVYVAVHFDVSHVDQEVINDSYSLLEFAYPTDRTREVLNDYA
KIMSHNVTAQVRRYFPETTE"
6759..7745
/locus\_tag="SAV0999"
6759..7745
/locus\_tag="SAV0999"
/codon\_start=1
/transl\_table=11
/product="similar to transcription factor"
/protein\_id="BAB57161.1"
/db\_xref="GI:14246769"
/translations="MLVALNEEKERVLTATLRKTOYPCVCGQVILKRGKLVISHF
AKHLABQKCFNNETIKHYKSLILAOIQOQCKVEIEPLKBIKOIPDILINNYY
IELOYSPITKQIILQRTGELKKMGYKSWLNDVDYCHNKVKFNHFSLSFINPTRKL
HTFNLQKQIMMFOIQYLGGHKYVAEKNAKIIIEFNEAPCDHAYVKLSKFAINQY
IKYCRWNSVLEPTLSAMYQLTDQEVVYNYGIFPEQIYIENHPIEWQLQVDLWLK
NGSKLVNDNLNYFKLKKFIVALESKTIIIEKLNNYLNICSDRGNVDQIILF"
7793..9601
/locus\_tag="SAV1000"
7793..9601
/locus\_tag="SAV1000"
/codon\_start=1
/transl\_table=11
/product="thimet oligopeptidase homologue"
/protein\_id="BAB57162.1"
/db\_xref="GI:14246770"
/translations="MSQQLSRREEQERKPEYTWDLTTIPKDDAEFAAPKFEVENEIGK
BEQFGHIGDSAEITLYNALIELEDTLTGLEKVVYVAHLKQDDTNDKTYGMESRAHQ
LIKFSWSAWFLVPEILIDEDKIOSFVNSYDKLOKAFDLKLINEKRPHIDAEATEK
LITEAQDALSTPSNYGFMNSADLFEADAIDKGNAPLTOCTFTIKYLSDDDRKLRES
AFRNYKYAGAHNNTLGATLAGEVKNVFNATHNYKTAREKALSNHHPENVYONLV
KTWHKYLPLLRHTELKELLGLDKMVDLYTPIKIDIKFEMPYEEAKEMWLKALEP
MGEEYLVNVVKEGLNRWVDYENKGRSGYSGAHLNPFILLNWSNTISDLYTLVH
EFGSHAHSYFSRKFQPSNSTDYTI FVAEVAETCNBALLSDYMDKHLDDKRLLLNQE
LERFRATLPRQTMFAEEFEKHIAIEAGEPLTPTRMNEEYAKLNKLYFGDSVETDEDI
LSERSRIPHFTYNNYVQYATGYSAAQSLSHOILLTEGKPAVDRIYINEFLKGGSNYP
EILKNAGVDMTTPTEPEQACEVFEQKLNAFEKLMKA"

gene		complement (10061. .10846)	
Alignment Scores:			
Pred. No.:	1.24e-42	Length:	347785
Score:	724.50	Matches:	149
Percent Similarity:	66.67%	Conservative:	63
Best Local Similarity:	46.86%	Mismatches:	95
Query Match:	42.87%	Indels:	11
DB:	1	Gaps:	4
US-10-724-972A-6352 (1-335) x AP003361 (1-347785)			
QY	13	ValIleGlyLeuLeuPheValLeuIleAlaThrAlaCysGlyAsnAsnSerSer	32
DB	46237	GTGTGTTTATCTAATCTTAGTTAGCAGTAGCGGTTGTGTCAAAAGATACT---	46293
QY	33	AsnSerSerLysGluSerSerLysAspGlyValGluIleLysHisGluGluGlyThr	52
DB	46294	-----GAAGAGAAACTGAATCAGCACATAAAGATGATTTAGGAACCTGAA	46341
QY	53	LysValProLysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeu	72
DB	46342	AAAAATTAAAGAAAATCCTAAACGTGTTGTATTAGAAATATAGTTTGTCTGATTATTA	46401
QY	73	ValAlaLeuAspValLysProValGlyIleAlaAspAsnLysLysAsnArgIleIle	92
DB	46402	GCAGCATTAGATAGAACTGTGTGTTATTCAGATGTCGACACTAAAAATATAACA	46461
QY	93	LysProLeuArgAspLysIleGlyLysThrThrSerValGlyThrArgLysGlnProAsn	112
DB	46462	AAGTCAGTAAGAGATAAGATTGGGGCATATGAATCGTTGGATCTAGACCGCAACCGAAT	46521
QY	113	LeuGluGluIleSerLysLysLysProAspLeuIleIleAlaAspAsnAsnArgHisLys	132
DB	46522	ATGGAGTGAATGAATTAATAACCGGATTTGATCATTCAGATGTTAGCAGACATAAG	46581
QY	133	GlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGly	152
DB	46582	AAAAATCAATCAGAAATTGACCAAAATGCTCCGACAATCATGTTAGTCGCGTACGGGA	46641
QY	153	AspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu	172
DB	46642	GATTATTAATGCAATATTGAAGCATTTAAACAGTCGCTAAAGCAGTAGGCGCAAGAGAAA	46701
QY	173	GluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGluTyrLysLysGluIle	192
DB	46702	GAAGCGCAGAGCGCTCGGAAAGCATGATAAAATATTAGCGGAGATTAGAAAGAAAAT	46761
QY	193	ThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAla	212
DB	46762	GAACAGAGTACGTTAAATCTGCATTTGCATTCGATCTCAAGAGCAGGTATGTTTATT	46821
QY	213	HisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu	232
DB	46822	AATAATGAGATACATTTATGGGACAACTCTTAATTAATGGGTATTTCAACCTGAAGTC	46881
QY	233	SerAspAspValThrLysGlyLeuSerLysTyrLysLysGlyProTyrLeuGlnMetAsn	252
DB	46882	ACAAAAGACAAAACCTACGCATGTTGTTGGAACCGCAAGCGGTGCTCTTATATATTAAAT	46941
QY	253	ThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn---LysAla	271
DB	46942	AATGAAGAATCTGCCAATATCAATCCAAAGATTATGATTTTATGACCACTGACGGAAACG	47001
QY	272	SerSerAsnGluProSerLeuLysGluLeuGluLysAspPro---ValTrpLysLysLeu	290
DB	47002	GACAAAATAGACGAAATTC-----ATTGATCTCGAGTTTGGAAATCATTA	47049
QY	291	AsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArg	310
DB	47050	AAAGCTGTGAAGATAACAAGTTTATGAGCTTGACCGAATAGTGGTTGAATCAAGG	47109
QY	311	GlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys	328

Db	47110	GGGATTATCGCAAGTGAAGATATGGCAGAAAGATTAGAAAAAATTGCAGAAAAA	47163
RESULT 14			
A64853/c			
LOCUS	A64853	2247 bp	DNA linear PAT 29-MAR-1999
DEFINITION	Sequence 10 from Patent WO9731114.		
ACCESSION	A64853		
VERSION	A64853.1	GI:4530844	
KEYWORDS	unidentified		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1		
AUTHORS	Burnham, M.K. and Hodgson, J.E.		
TITLE	POLYNUCLEOTIDES AND AMINOACID SEQUENCES FROM STAPHYLOCOCCUS AUREUS		
JOURNAL	Patent: WO 9731114-A 10 28-AUG-1997;		
FEATURES	SMITHKLINE BEECHAM PLC (GB)		
source	Location/Qualifiers		
	1..2247		
	/organism="unidentified"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32644"		
ORIGIN			
Alignment Scores:			
Pred. No.:	5.35e-45	Length:	2247
Score:	721.00	Matches:	169
Percent Similarity:	73.48%	Conservative:	36
Best Local Similarity:	60.57%	Mismatches:	69
Query Match:	42.66%	Indels:	9
DB:	6	Gaps:	1
US-10-724-972A-6352 (1-335) x A64853 (1-2247)			
QY	1	GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeu	20
Db	831	CGAGTGGGACNGATGAGAGGTCATAAAACTTTTAGTATATTGGGATTATAGTTCGCTTA	772
QY	21	IleAlaThrAlaAlaCysGlyAsnAsnSerSerSerAsn-SerSerLysGluSerSerly	40
Db	771	CTTTAGTTGCGAGCTTGTTGTAATACGGATAATTCAGTAAAGAGAAATCATCAACTAA	712
QY	40	sAspGlyVal-GluIleLysHisGluGluGlyThrThrLysValProLysHisProLysA	60
Db	711	AGATACTATTTCGGGTAAAGATGAAATGGTACAGTAAAGTACCTAAAGATGCACAA	652
QY	60	rgValValLeuGluTyrSerPheValAspAla---LeuValAlaLeuAspValLysP	79
Db	651	GTATCGTTGTATTAGAGTACTCATTTGCAAGGTTGCATTANCAGCATATAGCCGTTANAC	592
QY	79	ro-ValGlyIleAlaAspAsnLysLysAsnArgIleIleLysProLeuArgAspLys	98
Db	591	CAAGTTCGTATTCCTGATGATGTTAAGAAAAACGTTATCATAC-CCAGTAAAGAGAAAA	533
QY	99	IleGlyLysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLys	118
Db	532	TTGGGGGT-AATACTCTGTAGGTACACGTAA-CAAGCNAACTTAAAGAGANATTAGTAAA	475
QY	119	LeuLysProAspLeuIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeu	138
Db	474	TTAAACCCGGGATTAATATATCNCCTGATGACGATAGACATAGGTTATTAATAAGAAATTA	415
QY	139	AsnLysIleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIle	158
Db	414	GCCCATTCACACACA-TTATCATTTAAGAGTTTGTATGGAGCTACACCCCAATATT	356
QY	159	AspAlaPheLys-ThrIleSerLysAlaLeuGlyLysGluGluGlyLysArgLe	178
Db	355	AATTCGTTCAACAACAATTGCTAAAGCTTTAATAAAGANAAGAGCGGNAACGCTCT	296
QY	178	uGluGluHisAspLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnG1	198

Db	295	TGCTGAGCATGATAAAATTAATCAATAGTATAAAGATGAATTAATTAATTTGATAGAAATCA	236
Qy	198	nlysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTy	218
Db	235	CAAAAGTCTCCAGCAGTAGTCTGCTAAAGCTGCTTTTATAGCACATCCACACTATTCA	176
Qy	218	rValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrly	238
Db	175	TGTTGACAAATTTTAAACGANCCTAGGATTTAAAAATGCAATTAAGTGACGATGAACACA	116
Qy	238	sGlyLeuSerLysTyLeuLysGlyProTyLeuGlnMetAsnThrGluThrLeuSerGl	258
Db	115	AGGTTAAAGTAAATATTGAAAGGACCTTACTTACAATTAGACACCTGAACATTAGCTGA	56
Qy	258	nValAsnProGluArgMetPheLeuMetThrAsnLysAlaSerSerAsn	274
Db	55	TTTAAATCCAGAGCGTATGATCATTTATGACAGATCATGCTAAAAAAGAT	7
RESULT 15			
BX571856_10			
WPCOMMENT			
Sequence split into 29 fragments LOCUS BX571856 Accession BX571856			
Fragment Name Begin End			
BX571856_00 1 110000			
BX571856_01 100001 210000			
BX571856_02 200001 310000			
BX571856_03 300001 410000			
BX571856_04 400001 510000			
BX571856_05 500001 610000			
BX571856_06 600001 710000			
BX571856_07 700001 810000			
BX571856_08 800001 910000			
BX571856_09 900001 1010000			
BX571856_10 1000001 1110000			
BX571856_11 1100001 1210000			
BX571856_12 1200001 1310000			
BX571856_13 1300001 1410000			
BX571856_14 1400001 1510000			
BX571856_15 1500001 1610000			
BX571856_16 1600001 1710000			
BX571856_17 1700001 1810000			
BX571856_18 1800001 1910000			
BX571856_19 1900001 2010000			
BX571856_20 2000001 2110000			
BX571856_21 2100001 2210000			
BX571856_22 2200001 2310000			
BX571856_23 2300001 2410000			
BX571856_24 2400001 2510000			
BX571856_25 2500001 2610000			
BX571856_26 2600001 2710000			
BX571856_27 2700001 2810000			
BX571856_28 2800001 2902619			
Continuation (11 of 29) of BX571856 from base 1000001 (BX571856 Staphylococcus aureus su			
Alignment Scores:			
Pred. No.:	7.38e-43	Length:	110000
Score:	719.50	Matches:	147
Percent Similarity:	66.67%	Conservative:	65
Best Local Similarity:	46.23%	Mismatches:	95
Query Match:	42.57%	Indels:	11
DB:	1	Gaps:	4
US-10-724-972A-6352 (1-335) x BX571856_10 (1-110000)			
Qy	13	ValLeuGlyLeuPheValLeuLeuAlaThrAlaAlaCysGlyAsnAsnSerSer	32
Db	54238	GTGTGTTTATGCTATCTAGTTGTAGCAGTAGCGGTGCTGCTCAAAAGATACT---	54294
Qy	33	AsnSerSerLysGlySerLysAspGlyValGluLeuLysHisGluGluGlyThr	52
Db	54295	-----GAAGAGAAACTGAATGACGACGATAAAGATGAATAGGAACTGAA	54342
Qy	53	LysValProLysHisProLysArgValValValLeuGluTySerPheValAspAlaLeu	72

Db	54343	AAATTAAGAAAATCCGAAACGTATTGTTGATTAGAAATATAGTTTCTGATTATTTA	54402
Qy	73	ValAlaLeuAspValLysProValGlyLeuAlaAspAspAsnLysLysAsnArgIlelle	92
Db	54403	GCAGCATTAGATGAACCTGTTGGTATTGCGATGATGGTAGTACTAAAATATAACA	54462
Qy	93	LysProLeuArgAspLysIleGlyLysTyThrSerValGlyThrArgLysGlnProAsn	112
Db	54463	AAGTCTAGTAAAGATAAGATTGGGGCATATGAATCGTTGGATCTAGACCGCAACCGAAT	54522
Qy	113	LeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsnAsnArgHLeLys	132
Db	54523	ATGGAAGTGTATAGTAATTAATAACCGGATTGATTATTGCAGATGTCAGCAGACATAAG	54582
Qy	133	GlyIleTyLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGly	152
Db	54583	AAAAATCAAAATCAGAAATTTGCGCAAAATTTGCGCCGACCAATCATGTTAGTCGGTACGGGA	54642
Qy	153	AspTyAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu	172
Db	54643	GATTACAATGCAAAATATTGAAGCATTTAAACACAGTCCTAAACAGATTGGCAAGAGAAA	54702
Qy	173	GluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGluTyLysLysGluIle	192
Db	54703	GAAGCGAAAGAGCGTCTGGAAGAGCATATAAAATATTAGCGGAGATTAGAAAGAAAT	54762
Qy	193	ThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAla	212
Db	54763	GAACAGAGTAGCTTAAATCTGCATTTGCAATTTGCTATCTCAAGACGAGGTATGTTATT	54822
Qy	213	HisProSerAsnSerTyValGlyGlnPheLysSerGlnLeuGlyPheLysGluAlaLeu	232
Db	54823	AATAATGAAGATACATTTATGGGACAATTTCTTAATTAATAATGGGTATTCAACCTGAAGTC	54882
Qy	233	SerAspAspValThrLysGlyLeuSerLysTyLysLysGlyProTyLeuGlnMetAsn	252
Db	54883	ACAAAGACAAACTGCGCATCTTGGTGAACGCGAGGGTGGCCCTTATATTATTATAAT	54942
Qy	253	ThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn---	271
Db	54943	AATGAGAGCTTGCCAATATCAATCCAAAGTATTATGATTTTATGCTAGCAATGGAACAA	55002
Qy	272	SerSerAsnGluProSerLeuLysGluLeuGlyLysAspPro---ValTyLysLysLeu	290
Db	55003	GATAAAATAGAACGAAATTC-----ATTGATCTCGCATGTTTGGAAATCATTA	55050
Qy	291	AsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTyPalaArgSerArg	310
Db	55051	AAAGCTGTGAAGATAATAAAGTATATGATGTTGATCGAATAAGTGGTTCGAATCAAGA	55110
Qy	311	GlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys	328
Db	55111	GGTATTATGGCAAGTGAAGTATGGCAGAAGATTTAGAAAAAATTTGCAGAAAAA	55164

RESULT 16	AX951891/c	5718 bp	DNA	linear	PAT 08-JAN-2004
LOCUS	AX951891	Sequence 20 from Patent WO03095658.			
DEFINITION	AX951891				
ACCESSION	AX951891				
VERSION	AX951891.1	GI:40782280			
KEYWORDS					
SOURCE	Bacillus subtilis				
ORGANISM	Bacillus subtilis				
REFERENCE	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.				
AUTHORS	1				
TITLE	bj Rnvad,M.E., j Rgensen,P.L. and Hansen,P.K.				
JOURNAL	Homologous recombination into bacterium for the generation of				
FEATURES	polynucleotide libraries				
Source	Patent: WO 03095658-A 20 20-NOV-2003;				
	Novozymes A/S (DK)				
	Location/Qualifiers				
	1. .5718				

[illegible]





PMNLIITGKLPDDEGKSWKNVRVGYLDQHTVLEKGSIRDLVKDAFHFLPAMEEEMN  
EINYKMGCEADPDELEKLLBEVGVIODALTNDFYVIVDSKVEIARGLGSLDIGLERDV  
TDLGGQRTIYVLAKEKLLLEKPEYLLDDEPTNYLDQHLLEWLRKRYLQEVENAFILISHD  
IPFNSQNTIYVLAHVENQELTYVGDTHQFMVEYEVKKQLENAKKQOQVEAEKDFV  
ARNKARVSTNRMMASQKLDKMDMIELAAEKPEFFHFKPARTSGKLIETFKOLVIG  
YDSPLSPLRNMERGQKITALYANGIGKTIITLLKSLGEBIQPLEGSVERGSHIYTGFG  
EOEVEKTNNTCTIEEWSPEPSTQYVEIRAAPAKCOLATTKHIESRVSLSGGEKAKTV  
LCKLINSETNLLVLDDEPTNHLDAKEELKRALKEYKGSILLISHEDPYFMDIATETW  
NCESWITTKVL"  
1781. .2911  
/gene="yfmL"  
/locus\_tag="BSU07430"  
1781. .2911  
/gene="yfmL"  
/locus\_tag="BSU07430"  
/function="unknown"  
/note="similar to RNA helicase"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAB12572.1"  
/db\_xref="GI:2633067"  
/db\_xref="GOA:O34750"  
/db\_xref="UniProt/TREMBL:O34750"  
/translation="MTQTWPLFLHNAQSFIOENWNASGFKPTPVQEQAAQLIMDGKDV  
IAESTPTGKTLAVALPFLSRIPKEQHPQAVILASRELVMQIFQVIQDWKAGSELRL  
RASLIGGANVKQVEKLLKHPHIIIVGTPGVPELIIKAKLKMEHVXTIIVLDETQGLV  
PEHRTMQIITKLRLDQGLLCPSTLKKETEDVLRLEAQEPVLKVQSKAEAGKVK  
HQYLICDQRTIKVTLRQKLSRLGEMQALVFVRDIGNLSVYAEKLAHHVSLGVLSHSEAK  
KMERAKITATEDGEFFLLATDIAARGLDIENLPIYIHADIPEDEGYVHRSGRITGRA  
KNEGVLSTVLTKESKLLKMAKKLVGELSEAVYAGGUKTK"  
2979. .3425  
/gene="yfmK"  
/locus\_tag="BSU07440"  
2979. .3425  
/gene="yfmK"  
/locus\_tag="BSU07440"  
/function="unknown"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAB12573.1"  
/db\_xref="GI:2633068"  
/db\_xref="GOA:O34536"  
/db\_xref="UniProt/TREMBL:O34536"  
/translation="NASIDRFQWQOPELKILEHWFENEDTRRMGMILLLDEWYARV  
NKKDHUTVMAYDGLPAGMVIVFERBERTYIGLIVNPLRYLKGYGKQILQLKMTPEPD  
FTSVREWACIEEDNRISLAFQAGFTLEDEPDEDGLTLILRN"  
complement(3478. .4497)  
/gene="yfmJ"  
/locus\_tag="BSU07450"  
complement(3478. .4497)  
/gene="yfmJ"  
/locus\_tag="BSU07450"  
/function="unknown"  
/note="similar to quinone oxidoreductase"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAB12574.1"  
/db\_xref="GI:2633069"  
/db\_xref="GOA:O34812"  
/db\_xref="UniProt/TREMBL:O34812"  
/translation="MTASQQQLARRPQGIHVHDEFPTIPVEPKQGEVLVKTLYL  
VSDPYMRGRMDQTKSVSPVFLDIALSGVIAEYVSDGNKKGDIIVGNLSWQEFES  
AVSDPRLRGKDTSLASPASVAFDLGTMTGVTAYFGLLDIGRPKEGETVAVVSGAAGVGS  
TVGQIKAKGARVVGAGSDEKIDVLKQELQFDEAINYKTADDIQALQONACPDGVQV  
YFNVGGPITSDAVNNLLNEFARIPVCCGASISSYNAESEADMGPVRQSKLTKTSLMQG  
FIVSYDSDRFSEGAQLAEWLKAGKLHYEETIPDPAFLGLFKGENKKGQLILN  
VSDPS"  
complement(4979. .6199)  
/gene="yfmI"  
/locus\_tag="BSU07460"  
complement(4979. .6199)  
/gene="yfmI"  
/locus\_tag="BSU07460"





```

/transl_table=11
/product="hypothetical protein"
/protein_id="CAA74540.1"
/db_xref="GI:2226253"
/db_xref="GOA:O07616"
/db_xref="UniProt/TREMBL:O07616"
/translacion="MILIPCVTTLGHNHRNYQOERLEHMKTKLIIITLVLLISVLTAACSSSGNQNGKEHVAVTHDLKNTVPEHPKVWVLELGFIDTLDLGITPVGVAADNKAKOLINKVDLKIKDIGTVSGTRSQSMEKIASLKPDLLIADTTSHKKYVDOLKKIAPTIALNNINADYQDTIDASLTIAKAVGKEMEMKKLTAHEBKLSQTKQLSIANSQSULLIGTNDNIMARDENFFTRLTTOGVGYAIATSGNSDSDSGDSVNMMKMTLEQLLKTDPDVILLMTGTDDLDADGKRPIEKNVLWKLKAVKNGHYVHVDRVAVMSLRSDVGANAILDLEKEMPAAKK"
3099..3141
complement(3136..3731)
/gene="yhfr"
codon_start=1
complement(3136..3717)
/gene="yhfr"
/notes="similarity to a hypothetical protein o215b from E.coli (Swiss Prot P36942), and to several phosphoglycerate mutase (pgm) proteins, from Treponema pallidum (EMBL: U55214) and Schizosaccharomyces pombe (Swiss Prot P36623)."
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAA74541.1"
/db_xref="GI:2226254"
/db_xref="GOA:O07617"
/db_xref="UniProt/TREMBL:O07617"
/translacion="MTAVCLVRHGSETDNLOOKCQGKTDIPLNATGEROARETGXYVKDFSDIVTSPLKRAKRTAGETAINSYLHLPIVEMDDPKERYDGDAEGMPLEERTKRPVDNIYPNMTEBELTDLRMGLAEKALGNQVAPNNKVLIVAHGAHAHALLTEISGGDPFELASTRLVNACLNLIFBAEEKWRKDYNINSHLSGFIK"
complement(3722..3731)
/gene="yhfr"
complement(3788..4882)
/gene="yhfs"
complement(3788..4882)
/gene="yhfs"
/notes="similarity to acetyl-CoA:acetyltransferases (Thil, encoding the enzyme for the first step in poly-beta-hydroxybutyrate synthesis) from Chromatium vinosum (Swiss Prot P45369), Alcaligenes eutrophus (Swiss Prot P14611), Thiocystis violacea (Swiss Prot P45363), and B.subtilis (Swiss Prot P45855)."
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAA74542.1"
/db_xref="GI:2226255"
/db_xref="UniProt/TREMBL:O07618"
/translacion="MNNAVIVDAKRTIFGNQGLLKPFLPEDLAAPICRSLRKLEDQVDEVILNATGRGNLARLSALOGAPLSPVGMTIDRCQSGLEAVRYVASLIQAAGTMYTAGSESSSSPSFSRARFPDAIGPDGMIAEYTAARYSISRSMODEYALLSHQRSNRHADGEFYREEVVVALGELETDEAFULRPIEAIIIRPAKVPDETSSGTYVTANSSQIDAGAALLIMEEKAALKXPLFRFTGSVAGHPNPAPVPAIRQLGHTHDVTPDDIDLFEINEAPFAVKICVSQELGIPFKSINVRGGALGALGHYPGASGAAVTRLFVEYAKRRPCQYAAVAGSGGKIGLALLFEVLA"
complement(4879..6318)
/gene="yhfr"
complement(4879..6318)
/gene="yhfr"
/notes="similarity to long-chain-acyl-CoA synthetase from B subtilis (EMBL: 275208), bile acid-CoA ligase from Escherichia coli (Swiss Prot P19409), long-chain-fatty acid:CoA ligase from E.coli (Lcfa; Swiss Prot P29212)."
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAA74543.1"
/db_xref="GI:2226256"
/db_xref="GOA:O07618"

```

Db	2550	AATAATTTAAATGCTGATTATCAGAGTACAATTGACGCTTCGCTTACGATTGC	2609
Qy	168	LeuGlyLysGluGluGluGlyLysArgLeuGluGluHisAspLysLysLeuGlu	187
Db	2610	GTCGCAAGAGAGAAATCGAGAAAAGCTGACGCGCATGAAGAAAAGCTTAGCGAG	2669
Qy	188	TyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLys	207
Db	2670	ACAAAGACAAAATCAGCGCCAAC---AGCCAGTCCGTGCTTTTGATCGGAATACAAT	2726
Qy	208	SerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGly	227
Db	2727	GATACCATTTATGGCCAGGATGAAAACCTCTTTACATCGAGACTTTTAAACACAGGTGGC	2786
Qy	228	PheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGly---	246
Db	2787	TACCGATATGCAAATCAGT-----ACGTGAGGCAATACGGATTCAAGCAATGGCGGC	2837
Qy	247	-----ProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluAArgMet	264
Db	2838	GACTCTGTGAATATGAAAATGCACACTGGACGCTGCTGAAAACAGATCCGGATGTGATC	2897
Qy	265	PheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGlu---LeuGluLys	283
Db	2898	ATCTGTATGACAGGAAAAACAGATGACCTCGACCGCAGCGTAAACCCCGATCGAAAAG	2957
Qy	284	AspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg	303
Db	2958	AATGCTCTTTGGAAAAAATCGAAGCGAGTGA AAAACGGGCATGTATACCAACGATGGATCGT	3017
Qy	304	AspLeuTrpAlaArgSerArgGlyLeuIleSerSer-----GluGluMetAla	319
Db	3018	CGCGTGTGTCTCTCGCCCGCAGGTAGACGGGCGGAATGCCATTTTGGACGAGCTTCAA	3077
Qy	320	LysGluLeuValGluLeuSerLysLysAspSerLysLysAspAsn	334
Db	3078	AAAGAGATCCGGCTGCTAGAAATAAAGAAAAGAACGACGCAAC	3122

BSUB0006 198743 bp DNA linear BCT 07-JUL-2003  
 Bacillus subtilis complete genome (section 6 of 21): from 1011039  
 to 1209781.  
 Z99109 AL009126  
 Z99109.2 GI:32468723  
 Bacillus subtilis subsp. subtilis str. 168  
 Bacillus subtilis subsp. subtilis str. 168  
 Bacteria: Firmicutes; Bacilliales; Bacillaceae; Bacillus.  
 1 (bases 1 to 198743)  
 Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G.,  
 Azevedo, V., Bertolet, M.G., Beesieres, P., Bolotin, A., Borchert, S.,  
 Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S.C.,  
 Bron, S., Brouillet, S., Bruschi, C.V., Caldwell, B., Capuano, V.,  
 Carter, N.M., Choi, S.K., Codani, J.J., Connerthon, I.F., Cummings, N.J.,  
 Daniel, R.M., Denizot, F., Devine, K.M., Dusterhoft, A., Ehrlich, S.D.,  
 Emmerson, P.T., Enkian, K.D., Errington, J., Fabret, C., Ferrari, E.,  
 Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Gallizzi, A.,  
 Galleron, N., Ghim, S.Y., Glaeser, P., Goffeau, A., Golightly, E.J.,  
 Grandi, G., Guiseppi, G., Guy, B.J., Haga, K., Haiech, J., Harwood, C.R.,  
 Henaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hulio, M.F.,  
 Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y.,  
 Klaer-Blanchard, M., Klein, C., Kobayashi, Y., Koetter, P.,  
 Koningstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A.,  
 Lardinois, S., Lauber, J., Lazarevic, V., Lee, S.M., Levine, A., Liu, H.,  
 Masuda, S., Mausel, C., Medigue, C., Medina, N., Mellado, R.P.,  
 Mizuno, M., Moestl, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M.,  
 Ogawa, K., Ogihara, A., Oudega, B., Park, S.H., Parro, V., Pohl, T.M.,  
 Portetelle, D., Porwollik, S., Prescott, A.M., Presecan, E., Pujic, P.,  
 Purnelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M.,  
 Rivolta, C., Rocha, E., Roche, B., Rose, M., Sadaie, Y., Sato, T.,  
 Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J.,

TITLE	The complete genome sequence of the gram-positive bacterium <i>Bacillus subtilis</i>
JOURNAL MEDLINE PUBLISHED	Nature 390 (6657), 249-256 (1997) 98044033 9384377
REFERENCE	2 (bases 1 to 198743)
AUTHORS	Kunst, F., Ogasawara, N., Yoshikawa, H. and Danchin, A.
TITLE	Direct Submission
JOURNAL	Submitted (27-JUN-2003) I. Moszer, A. Danchin, Institut Pasteur, Genetique des Genomes Bacteriens, 28 rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr, adanchin@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45 68 89 48
COMMENT	On Jul 7, 2003 this sequence version replaced gi:2633260. This entry contains data from release R16.1 of the Subtilist database. Further data on gene annotation and detailed information about changes from previous releases can be found at <a href="http://genolist.pasteur.fr/Subtilist/">http://genolist.pasteur.fr/Subtilist/</a> .
FEATURES	Location/Qualifiers
source	1. 198743 /organism="Bacillus subtilis subsp. subtilis str. 168" /mol_type="genomic DNA" /strain="168" /db_xref="taxon:224308" complement (38. 1504) /gene="lytP" /locus_tag="BSU09370" complement (38. 1504) /gene="lytP" /locus_tag="BSU09370" /function="cell wall lytic activity" /note="alternate gene name: cwIE, yhdP" /codon_start=1 /transl_table=11 /product="gamma-D-glutamate-meso-diaminopimelate muropetidase (major autolysin) (CWBP49) " /protein_id="CAB12776.1" /db_xref="GI:2633272" /db_xref="GOA:O07532" /db_xref="UniProt/TrEMBL:O07532" /translation="MKKKLAAFTASAIIVGTTLVVTPAEAAITIKVSGDSLKWLQAOTY NLSVAALTSAHLSTGLVLSIGQTLTIPGSKSTSTSTSTSTMTKSGGSVTVKSGDSL WLIANEFKMTQVELKGLNGSSDLIRAGQKLKVGTVSSSSSSSKSNKNSKSSSSSS SSSNKSSSSSSSTGYKVLQGLDLSWKIANKVNMSIAELVLNNLKSDTIYVNVQLKTK SSGSDTSSKSNKSNSTSTSTTYKTVKSGDSLWKIANNNTLTVQIIRNNLNKSDVLY VGQVLKLTGRASS NNLTKDVLQGVQKLVIGKASS TTPSGFDCSGFYIYVLANKVTSVSLTAAGYNTWKSVSQPAVGDFVFSTYKAGPSHV GIYLGNGEFINANDSGVVISNMNNSYKQRYLGAKRYP" complement (1657. 2097) /gene="yhdP" /locus_tag="BSU09380" complement (1657. 2097) /gene="yhdP" /locus_tag="BSU09380" /function="unknown" /note="similar to unknown proteins" /codon_start=1 /transl_table=11 /protein_id="CAB12777.1" /db_xref="GI:2633273" /db_xref="GOA:O07573" /db_xref="UniProt/Swiss-Prot:O07573" /translation="MKLTNTDYSRLVLIIFLAERPGELSNIKOIAETYSISKNNLHK VIVRIYLGQVYVETIRGRCGTRICMPDENITGRVATKTRDENITVECEANQKCVI

[illegible]

Qy	208	SerGlyLeuLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGly	227
Db	96593	GATACCATTTATGGCCAGGGATGAATACTCTTTACATCCAGACTTTTAAACAGAGTCGGC	96652
Qy	228	PheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGly---	246
Db	96653	TACCGATATGCAATCAGT-----ACGTCAAGCAATAGCGAATTCAAGCAATGGCGGC	96703
Qy	247	-----ProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMet	264
Db	96704	GACTCTGTGAATAGAAATGACACTGGAGCAGCTGCTGAAACACAGATCCGGATGTGATC	96763
Qy	265	PheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGlu---	283
Db	96764	ATCCTGATGACAGGAAAAACAGATGACCTTCAGCGCGCAGCGTAAACGCCCATCGAAAAG	96823
Qy	284	AspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg	303
Db	96824	AATGTCCTTTGGAAAAAACTGAAGGCGAGTGAATAAACGGGCATGTATATACCGATGGATCGT	96883
Qy	304	AspLeuTrpAlaArgSerArgGlyLeuIleSerSer-----GluGluMetAla	319
Db	96884	GCGGTGTGGTCTCTGGCGCGCAGTGTAGACGGGCGCAATGCCATTTTGGACGAGCTTCAA	96943
Qy	320	LysGluLeuValGluLeuSerLysLysAspSerLysLysAspAsn	334
Db	96944	AAAGAGATCCGGCTGCTGAAGAAATAAAGAAAGACAGCGCAAC	96988

RESULT 21  
 ER017332 11

RESULT 21  
AE017333 11  
WPCOMMENT

Sequence split into 43 fragments LOCUS AE017333 Accession AE017333

Fragment Name	Begin	End
AE017333_00	1	110000
AE017333_01	100001	210000
AE017333_02	200001	310000
AE017333_03	300001	410000
AE017333_04	400001	510000
AE017333_05	500001	610000
AE017333_06	600001	710000
AE017333_07	700001	810000
AE017333_08	800001	910000
AE017333_09	900001	1010000
AE017333_10	1000001	1110000
AE017333_11	1100001	1210000
AE017333_12	1200001	1310000
AE017333_13	1300001	1410000
AE017333_14	1400001	1510000
AE017333_15	1500001	1610000
AE017333_16	1600001	1710000
AE017333_17	1700001	1810000
AE017333_18	1800001	1910000
AE017333_19	1900001	2010000
AE017333_20	2000001	2110000
AE017333_21	2100001	2210000
AE017333_22	2200001	2310000
AE017333_23	2300001	2410000
AE017333_24	2400001	2510000
AE017333_25	2500001	2610000
AE017333_26	2600001	2710000
AE017333_27	2700001	2810000
AE017333_28	2800001	2910000
AE017333_29	2900001	3010000
AE017333_30	3000001	3110000
AE017333_31	3100001	3210000
AE017333_32	3200001	3310000
AE017333_33	3300001	3410000
AE017333_34	3400001	3510000
AE017333_35	3500001	3610000
AE017333_36	3600001	3710000
AE017333_37	3700001	3810000
AE017333_38	3800001	3910000
AE017333_39	3900001	4010000

AE017333_40	4000001	4110000	
AE017333_41	4100001	4210000	
AE017333_42	4200001	4222645	
Continuation (12 of 43) of AE017333 from base 1100001 (AE017333 Bacillus lichenii)			
Alignment Scores:			
Pred. No.:	1 64e-34	Length:	110000
Score:	608.50	Matches:	132
Percent Similarity:	61.66%	Conservative:	69
Best Local Similarity:	40.49%	Mismatches:	118
Query Match:	36.01%	Indels:	7
DB:	1	Gaps:	5

US-10-724-972A-6352 (1-335) x AE017333 11 (1-110000)

QY	9	LysIleLeuSerValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCysGlyAsn	28
DB	11532	AAATATTTTGGCTATTTTATTTTATTTATGATGATATCAGTGATGACGCGCTCTCGGA--	11588
QY	29	AsnSerSerSerAsnSerSerLysGluSerSerLysAspGlyValGluIleLysHisGlu	48
DB	11589	-----GGNAGCGAGACGACGAAACTTCGGGTGAGATTCGCGTACAGTAAAGCATGAT	11642
QY	49	GluGlyThrThrLysValProLysHisProLysArgValValValLeuGluTyrSerPhe	68
DB	11643	CTTGAGAAACGAAAGTGGCGAATAACCGAAAAAGATTGGTCTTGATTTGGGATTT	11702
QY	69	ValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsnLysLys	88
DB	11703	ATTGATGGCTCTTTCATCGCGGTATTAAGCCGTGCGAATTTCGACGACGCGCAACCT	11762
QY	89	AsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArg	108
DB	11763	AAGTTTATTACGAGAAGTCAGGGGAAAAATCAAAAGGGTATATCTTCACTCGGTTCGCGC	11822
QY	109	LysGlnProAsnLeuGluGluIleSerLysLysProAspLeuIleIleAlaAspAsn	128
DB	11823	GCCAGCCCAAGCTTTGAAAAAATTCCTCTTTAAAGCCGATTATTAATATGCCGATTCG	11882
QY	129	AsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLys	148
DB	11883	AGCAGGCACGCGCGCTATGATAAGCTGTGAAAAATTGCGCCGACAATCGCGCTCAAA	11942
QY	149	SerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeu	168
DB	11943	AAC TTGAATGCGGATTATCAGCACACGCTTGATCATCTCTTACGATTGCGAAAGCCGTC	12002
QY	169	GlyLysGluGluGluGlyLysArgLeuGluHisAspLysLysIleGluGluTyr	188
DB	12003	GGCAAGAAAGCGCAATGGAGAAAAAATTGGCTGAACATAACAAAGCTGGATGATG	12062
QY	189	LysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSer	208
DB	12063	AAACGAAAAATTCGCG--AGCGGAAACAGACGATCTCTTCCTCGGGAACACAAATGAA	12119
QY	209	GlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPhe	228
DB	12120	GAATACCCGTCGCGATGAAACTTTTTCAGCTCTCAGCTTCTGACGAAATCCGCTAT	12179
QY	229	LysGluAlaLeuSerAspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyr	248
DB	12180	ACATACGCTGTCGGAGACAGC--GGCAAGGCGATCGGAAAAACGGTGAATCCGTCAAT	12236
QY	249	LeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThr	268
DB	12237	ATTAAATGACGTCGAACAGCTGCTTGAAGAAAGATCCCGACGTTATTGTCTCATGACG	12296
QY	269	--AsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGluLysAspProValTrp	287
DB	12297	GGAGAAAAAGATAAGTCGATGAAGACGGGAAAGACCGATCGAAAAAAGATCCTCTTTGG	12356
QY	288	LysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAla	307

D <sub>b</sub>	12357	AAAAGCTCAGCGCAGTCAAAAAACGGCAGGTTTATGAAGCCGCAGATTTCGCTGGTCTG	12419
Q <sub>y</sub>	308	ArgSerArgGlyLeuIleSerSerGluGluMetalLysGluLeuValGlu--LeuSer	326
D <sub>b</sub>	12417	CTCCGACCGCAGCATGTACCGCAGCAGTAGCTGATGGATGAATCGATCAGCATCTTCTG	12476
Q <sub>y</sub>	327	LysLysAspSerLysLys	332
D <sub>b</sub>	12477	TCTGAGGATCAAAAGAAA	12494

## RESULT 22

CP000002\_11

Sequence split into 43 fragments LOCUS CP000002 Accession CP000002

Fragment Name	Begin	End
CP000002_00	1	110000
CP000002_01	100001	210000
CP000002_02	200001	310000
CP000002_03	300001	410000
CP000002_04	400001	510000
CP000002_05	500001	610000
CP000002_06	600001	710000
CP000002_07	700001	810000
CP000002_08	800001	910000
CP000002_09	900001	1010000
CP000002_10	1000001	1100000
CP000002_11	1100001	1210000
CP000002_12	1200001	1310000
CP000002_13	1300001	1410000
CP000002_14	1400001	1510000
CP000002_15	1500001	1610000
CP000002_16	1600001	1710000
CP000002_17	1700001	1810000
CP000002_18	1800001	1910000
CP000002_19	1900001	2010000
CP000002_20	2000001	2110000
CP000002_21	2100001	2210000
CP000002_22	2200001	2310000
CP000002_23	2300001	2410000
CP000002_24	2400001	2510000
CP000002_25	2500001	2610000
CP000002_26	2600001	2710000
CP000002_27	2700001	2810000
CP000002_28	2800001	2910000
CP000002_29	2900001	3010000
CP000002_30	3000001	3110000
CP000002_31	3100001	3210000
CP000002_32	3200001	3310000
CP000002_33	3300001	3410000
CP000002_34	3400001	3510000
CP000002_35	3500001	3610000
CP000002_36	3600001	3710000
CP000002_37	3700001	3810000
CP000002_38	3800001	3910000
CP000002_39	3900001	4010000
CP000002_40	4000001	4110000
CP000002_41	4100001	4210000
CP000002_42	4200001	422233

SOURCE  
ORGANISM Pasteurella multocida subsp. multocida str. Pm70  
Pasteurella multocida subsp. multocida str. Pm70  
Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
Pasteurellaceae; Pasteurella.  
REFERENCE  
AUTHORS May,B.J., Zhang,Q., Li,L.L., Paustian,M.L., Whittam,T.S. and Kapur,V.  
TITLE Complete genomic sequence of Pasteurella multocida, Pm70  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (6), 3460-3465 (2001)  
MEDLINE 21145866  
PUBMED 11248100  
REFERENCE  
AUTHORS Zhang,Q. and Kapur,V.  
TITLE Direct Submission  
JOURNAL Submitted (24-OCT-2000) Department of Veterinary Pathobiology,  
University of Minnesota, 1971 Commonwealth Ave., St. Paul, MN  
55108, USA

FEATURES  
source  
1. .13792  
/organism="Pasteurella multocida subsp. multocida str.  
Pm70"  
/mol\_type="genomic DNA"  
/strain="PM70"  
/sub\_species="multocida"  
/db\_xref="taxon:272843"  
87. .809  
/gene="PM0116"  
87. .809  
/gene="PM0116"  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AAK02200.1"  
/db\_xref="GI:12720329"  
/translation="MTTFIAGNINDYPLDSIPRELLIEKLRQSSNQVRKHRC  
RVAFHLLQLLKI SQKPTALLKHIDYSQSGRQLPVHDVDNISHGWDVAIIIRVN  
TQDGSIVGIDIESPKERYPALLAYFASQAWPTQQOEKEKSFYLSWCLREAVLK  
SQGVIVLUSAVQHDPLLRQIQSHCPAGQLVFSHRLPFYLAFFINOTCLDHYFTW  
QGGILAEITLDAIYYKNEKT"  
994. .2253  
/gene="hflk"  
/note="synonym: PM0117"  
994. .2253  
/gene="hflk"  
/codon\_start=1  
/transl\_table=11  
/product="hflk"  
/protein\_id="AAK02201.1"  
/db\_xref="GI:12720330"  
/translation="MENVMSLNGSDQDPWPKPGKKPESQRPDNEGQPNWINDNOSGR  
KONNOEQSPDDEIEIFNNLLKLGSGNGKRNNSHSGSLNGLKLLPIAAVIGAIW  
GVSGFYTKAEARGVVRFGELHAI VQPLAWKPTFIDRVI PVNVEQVKELRTQGSML  
TODENMKVEMTVQVRHDPKAYLFSTVNADDSLNQATDSALRYVIGHMSDDILTTG  
RSVRENTWKTLNTLIIIEPYNMGLEVVDVQSPARPEVKDAFDDAIKAEDEQRYIR  
BAEYAREREPARIADQAQRILEEATAYKDRVYLDKAGEVERFLLPFEKAPPELLRE  
RLYIQTMEKVMPTKPMVDGNSGNNLTVLPLEQILKQQTKGKSTNATSEVRPVFR  
AETPNMPTQTSSTSTQENTRQGRFN"  
2253. .3140  
/gene="hflc"  
/note="synonym: PM0118"  
2253. .3140  
/gene="hflc"  
/codon\_start=1  
/transl\_table=11  
/product="hflc"  
/protein\_id="AAK02202.1"  
/db\_xref="GI:12720331"  
/translation="MRKFLTEVIVVIAALYSSIVISVSEGRGIMLRFKSKVHRDANK  
VVVYNPGLHFIPIDSIKILDARDITLDGADRFTVVEKKDLILDVSVKWRISDFGR  
FYATGGDYAQASNLLRRKYNDRIRSETSGRTIKDIVSGTRGELMGAKALNTGPR  
STAEGLIEVDVVRKQINLPDEVSSSIYQNRRAERDAVAREHRSQGRKAAFIQADV  
RKVTLIIANARNTAQELRGSDATAAKVFSDAFSQEPQFVSFLRSLKAYESSFANSON  
MMILKPSDSDFRFRMQAPTK"

gene  
complement (3197. .3739)  
/gene="PM0119"  
complement (3197. .3739)  
/gene="PM0119"  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AAK02203.1"  
/db\_xref="GI:12720332"  
/translation="MAKRGKNKIEWNTWEOEEBIIIWVSKSEIKRDAEALKKLGK  
LVLDTKAKLEKIPIEKKLELEAQLRQKREARRQLOVIGKLLRNIDIVEPIQDALEK  
LENKQQQQAMHLKLEMLRDLTLDNPEAITQFLEQHPDADRHRLNLRVAAQKEKAQN  
KPPKAYREIFQYKDFMLED"  
3834. .5189  
/gene="pmba"  
/note="synonym: PM0120"  
3834. .5189  
/gene="pmba"  
/codon\_start=1  
/transl\_table=11  
/product="Pmba"  
/protein\_id="AAK02204.1"  
/db\_xref="GI:12720333"  
/translation="MONTTEKSTALLKQOEOTLREAVRYAVDLAQKAGATAEVAVTKVS  
GLSVSTRLOEIBNVEFTNDGALGISVILGNOKGNASTSDLSPEAIKNTVYEAASIAKY  
ASPDSCAGLADYDLMAFDAPDLALYHAADIDVEKATATALETEQAALSYDKRIVNSEG  
SAPNSHTGVRVYCNTHGMLQSYLSRYSLSVCSVIAGVEDQLERDEYTTIRDLNQLES  
ASWGEHCAKTLTSLRQPKMATQVPIFLNDVATGLISHLTAAISGSLYRKSFL  
LDHLGKQVLPMFPHISQPHLIGQLASSPFDFSEGVRTQCREIIQQGTQTLTLYTSYG  
KKLGNRSTHAGGIHNLVKNPTTGLTTLTLLROMGTGLLVTDLMGCVGNLVNVTGYSRG  
AAGFWVENGEIQYPAVEITIAQOLKEMLRNIVAGDDIIEHRSNIQTGSLLEHLKVS  
N"  
5424. .5963  
/gene="hpt"  
/note="synonym: PM0121"  
5424. .5963  
/gene="hpt"  
/codon\_start=1  
/transl\_table=11  
/product="Hpt"  
/protein\_id="AAK02205.1"  
/db\_xref="GI:12720334"  
/translation="MKKHVGILISEQDVRARIOTLGREITQYVQKAVEKIVVVGLL  
RGSEFMADLVRELNLVPEIEFMTTSSYSGSGMTTNHDVKISKDLGDIKGBHVLIED  
IIDTGYTLQKREILKUREPALSRLICTLLDKPSRREVEVPVVEWVGFSIPDFVVGYGI  
DYAQRYNNGYIGHVLEE"  
5996. .6580  
/gene="gmha"  
/note="synonym: PM0122"  
5996. .6580  
/gene="gmha"  
/codon\_start=1  
/transl\_table=11  
/product="Gmha"  
/protein\_id="AAK02206.1"  
/db\_xref="GI:12720335"  
/translation="MYLEQIKAELENAATVNLKFIENDENNMKLIQQAALLISDSFKQG  
KVLSCGNGSCHDMFAEELTGYRENRPYPAIAISDVSHLSVCVSNDRPGYVFS  
RYVEAVQKGDVLFGLSTSGNSKNILNAIKAAKEKMKVIAMTCKDGGQAGIADVEI  
RVPHFGVADRIOEIHKVILHMLLIEFEAKTE"  
6756. .7487  
/gene="artp"  
/note="synonym: PM0123"  
6756. .7487  
/gene="artp"  
/codon\_start=1  
/transl\_table=11  
/product="Artp"  
/protein\_id="AAK02207.1"  
/db\_xref="GI:12720336"  
/translation="MAIRVKNLNFYSSQALFDINLSAEGDTVLLIGPSGACKSTL  
IRTLNLLEVPASGELIIANSQFDLSETKTPKQICRLRQDVGVFQYHLWPHLTVIE  
NLIEAPKVKVLKLSDEQARSALBLRLRLAQADRFPLHLSGGQQQRVAIARALMMK





Query Match:	30.21%	Indels:	3
DB:	6	Gaps:	3
US-10-724-972A-6352 (1-335) x AX433651 (1-801)			
Qy	60	ArgValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysPro	79
Dy	3	AAAGTTGGTCTTGAATGGGATTTATTGATCGCTCTCTGATCGGATTAAGCCT	62
Qy	80	ValGlyIleAlaAspAsnLysLeuAsnArgIleIleLysProLeuArgAspLysIle	99
Dy	63	GTCCGATTCGGAGCAGCGCAACCTTAAGTTTATTACGAGAGGTTCAGGGGAAAATC	122
Qy	100	GlyLysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeu	119
Dy	123	AAAGGGTATATCTTCAGTCGGTTCGCGCCAGCCAAAGCTTTGAAAAAATGCTCTTTA	182
Qy	120	LysProAspLeuIleAlaAspAsnAsnArgHisLysGlyLeTyrLysAspLeuAsn	139
Dy	183	AAGCCGATTTAATTATTCGCGATTCGAGCAGCAGCAGCGGCTCTATGATAGCTGTCG	242
Qy	140	LysIleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAsp	159
Dy	243	AAATTCGCGCGAATCGGCTCAAAAACCTTGAATGCGCGATTCAGGACGCTTGAT	302
Qy	160	AlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGluGlyLysLysArgLeuGlu	179
Dy	303	GCATCTCTACGATTCGGAAGCGTCGCAAGAAGCGCAATGCGAGAAAATTTGGCT	362
Qy	180	GluHisAspLysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLys	199
Dy	363	GAACATAAACAAAGCTGGATGAATTCAAACAGAAATTCGGC---AGCGGAACAGAGC	419
Qy	200	ValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrVal	219
Dy	420	ATTCCTCTCGTCCGGACACAAATGAAGAAATCACCGTCGCGATGAAAACCTTTTTCAG	479
Qy	220	GlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGly	239
Dy	480	TCTCAGCTTCTACGAAATCGCTATACATCGGTGTCGGAGACAGC---GGCAAGGC	536
Qy	240	LeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnVal	259
Dy	537	GATCGGAAACGGTGAATCCGCTCAATATTAATGACCTCGACAGCTGCTTGAGAA	596
Qy	260	AsnProGluArgMetPheIleMetThr---AsnLysAlaSerSerAsnGluProSerLeu	278
Dy	597	GATCCGACGTTATTGTCCTGATGACGGAGAGAAAGATAAAGTCGATGACGACGGA	656
Qy	279	LysGluLeuGluLysAspProValTyrLysLysLeuAsnAlaValLysAsnGlnArgVal	298
Dy	657	AGACCATCGAAAAAGATTCCTTTTGGAAAAAGCTCAGCGCAGTCAAAAACGGCA	716
Qy	299	AspIleLeuAspArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMet	318
Dy	717	TATGAAGCGCAGATTCGCTCGTCCGACCGCAGCATTCGCGAGCAGATGAGCTG	776
Qy	319	AlaLysGluLeu 322	
Dy	777	ATGATGAATC 788	

RESULT 25  
AB082122  
LOCUS  
DEFINITION  
Vibrio parahaemolyticus ferric vibrioferrin transport system gene cluster (pvuB, pvuC, pvuD, pvuE), complete cds.  
AB082122  
AB082122.1 GI:23307109  
KEYWORDS  
vibrio parahaemolyticus  
vibrio parahaemolyticus  
Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio.

REFERENCE AUTHORS TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE JOURNAL	1 Tanabe,T., Funahashi,T., Nakao,H., Miyoshi,S., Shinoda,S. and Yamamoto,S. Identification and characterization of genes required for biosynthesis and transport of the siderophore vibrioferrin in Vibrio parahaemolyticus J. Bacteriol. 185 (23), 6938-6949 (2003) 22982453 2 (bases 1 to 4044) Yamamoto,S., Tanabe,T. and Funahashi,T. Direct Submission Submitted (22-MAR-2002) Shigeo Yamamoto, Okayama University, Faculty of Pharmaceutical Sciences; 1-1-1 Teushima-naka, Okayama, Okayama 700-8530, Japan (E-mail:syamamoto@phs.u-oc.jp, Tel:81-86-251-8473, Fax:81-86-251-7926) Location/Qualifiers 1. .4044 /organism="Vibrio parahaemolyticus" /mol_type="genomic DNA" /strain="WPI" /db_xref="taxon:670" 156. .1079 /gene="pvuB" 156. .1079 /gene="pvuB" /codon_start=1 /transl_table=11 /product="ferric vibrioferrin-binding periplasmic protein PvuB" /protein_id="BAC16540.1" /db_xref="GI:23307110" /translation="MNKKQFTIKHFTVLLMLSSVPMSPSALSQAQSVQEQGTFE LEAPQRIWLEFSDVADNDIAIRVPAVKIEPMQSVGMSQP SLEATVUKPDIIADAEHRRAIYQDLQRIAPTLLKSGEYQENLSAQIKGVAIG KQAMQRIELHROTAEFKQHFATQETIQGVSDKGMWLSHRSYVAGGVSTLIGIQ SPLAPSEQNAYIPTSFEKLLNPDMLLVGLYSQNPVDEWRKNFLFKLLTAAKKRL VEVSPSLMNLARKMLAAEIAARNLEALLGRS" 1076. .2113 /gene="pvuC" 1076. .2113 /gene="pvuC" /codon_start=1 /transl_table=11 /product="ferric vibrioferrin transport system permease protein PvuC" /protein_id="BAC16541.1" /db_xref="GI:23307111" /translation="MSASIISESPTSVRKPHWKASLFVCAMALCTLSCYTASMGWSN FLSVNDLTSYFAPDEGNMLHQILATLRAPRTYAGLLIGASLAVSGVLMOGLTRNPL APSILGINAGAACFMALASTGVPFPSQLNPIINAVFALLSGGAVMLGGFFBSRSH PLRLVLAGIALSALLIGETRASVLIADDMAYSVLHWTLSLSAVSDSQMQLWPFATL GLVLAMGLARNLLAGDEAVGGLGNIRITRISGLAVVLLAGTSVAIGPIGFVG LLVPHLVPRPIVGHNYHILI PVSALCGAALVTSDALSRATFAETPVGVITALLGTP CFVIAMRKSS" 2110. .3081 /gene="pvuD" 2110. .3081 /gene="pvuD" /codon_start=1 /transl_table=11 /product="ferric vibrioferrin transport system permease protein PvuD" /protein_id="BAC16542.1" /db_xref="GI:23307112" /translation="MMNQIKMALIGLLIASSITLFGAANLSAQOVFALLFSFSDS DFVHQVLRPMLLAIGVAGLISGLVQVIRNPILASPDLMGISAGAGLAATACLV LYPNAVMLPMVAMAGGLAACGFIADVYMSKPTPARLALIGVASFVAGSIDFLL IVHPIENTAMVLTGSLGRNNQQVPIWSALLLLPLAFWLAWRLDFGNHKLITPASA TLGKPKQIILALIAAVLLASISVSAGTISFVGLLAPHLARLLFGHNHKLITPASA TLGALLVICADGLARGLQPPIELPAGVLTISVIGAPYIFLLYRGN" 3084. .3848
gene	
CDS	
gene	
CDS	
gene	
CDS	
gene	

```

CDS
/gene="pvuE"
3084..3848
/gene="pvuE"
/codon_start=1
/transl_table=11
/product="ferric vibrioferrin transport ATP-binding
protein PvuE"
/protein_id="BAC16543.1"
/db_xref="GI:23307113"
/tranlations="MLRTQNLVAYGKQTIIPDLISIPGKTIALVGNPGCKSTILL
KTLVRLNKPASGVLDYDKPLSSYGDKRLARSLFVQILVSEGITVRKLVBGRSP
YLSHMRKLSQDQALVEQAMQDVTGVEFADKPVESLSGQQRQAWTAMVAQDTDVM
LDPEYTLDSLHQVELMKLRKMDKGTVVVVLHDLNQACRYCDHLVLEKQLVTO
GTQEVLTGKMLASVFDLQARVNDPISGTPMCVAV"
3897..3937
terminator
ORIGIN
Alignment Scores:
Pred. No.: 1.78e-27 Length: 4044
Score: 492.00 Matches: 113
Percent Similarity: 55.11% Conservative: 65
Best Local Similarity: 34.98% Mismatches: 103
Query Match: 29.11% Indels: 42
DB: 1 Gaps: 5

US-10-724-972A-6352 (1-335) x AB082122 (1-4044)
QY 31 SerSerAsnSerSerLysGluSerSerLysAspGlyValGluIleLysHis----- 47
DB 138 AGCACTGGAAGAACGACATGAAATAATAAATAAACAATTACGATCAACACTTTTCAACC 197
QY 47 ----- 47
DB 198 GTTTTGCTGATGCTGCTGAGCTCAGTCTTTTATGTCATTTAGCGCATTTCTCAAGCGCGC 257
QY 48 -----GluGluGlyThrThrLysValProLysHisProLysArgValValVal 63
DB 258 AGCGTGCAAGATGACCAAGGCACCTTTTGAGCTAGAGCATACCTCAACGATTTGTGTT 317
QY 64 LeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAla 83
DB 318 TTGGAGTTTCTGTTGTGATGCTCTGGCGCGCGTGTCTCACCAGTAGCGGTGCTGCC 377
QY 84 AspAspAsnLysLysAsnArgIleLysProLeuArgAspLysIleGlyLysThr 103
DB 378 GATGACAAACGATGCCATACGAGTATCCTCTGCTGTGCGCAAGATTCGAGTCTCAAGCTGATTG 497
QY 104 SerValGlyThrArgLysGlnProAsnLeuGluLysSerLysLysLysProAspLeu 123
DB 438 TCGGTGCGCATGCGCTCACAAACCGAGTTTAGAGCGATTGCGGTCTCAAGCTGATTG 497
QY 124 IleIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaPro 143
DB 498 ATCATTCGCGATGCCAAGCTCATCGCGGATTTATCAAGATTTGACGCGAATTGCGCA 557
QY 144 ThrIleGluLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThr 163
DB 558 ACTTTGCTGCTAAAAGCGGTGGAGACATATCAAGAAAATTCGAATCCGACACAAA 617
QY 164 IleSerLysAlaLeuGlyLysGluGluGlyLysArgLeuGluLysAspLys 183
DB 618 ATCGCGGTGCGAATTGCGAAGCAAGCGCAATGACCGAAGATTCAGCTGCACACAG 677
QY 184 LysIleGluGlyLysLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAla 203
DB 678 ACCATCGCGAGTTCAACAGCATTTTTCGCG-----ACACAAGAACCAATTCAGTTTGGT 731
QY 204 ValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeu 223
DB 732 GTGTTTCCGATAAAGGATGTGCTGCATAGCGGGTTTCTTACGCTGCTGCGGTTTG 791
QY 224 SerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyr 243

```

```

DB 792 AGTACACTAGGCATCCAAAGCCCTCTGGCTCGTCTGAG----- 830
QY 244 LeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263
DB 831 ---CAAAATGCTTATATCCCAACAGCTTCGAACTTTTACTAAACCAACCTGATTGG 887
QY 264 MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeu---LysGluLeuGlu 282
DB 888 TTTACTAGTT-----GGACTTTTACTCTCAACCAATATCTGATGATGATGGCGC 935
QY 283 LysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAsp 302
DB 936 AAAAAATCCATTATTTAAGCTGTTAACTGGCGGCAAAAGCGACAACCTGGTTCGAAGTCT 995
QY 303 ArgAspLeuTrpAlaArgSerGlyLeuIleSerSerGluGluMetAlaLysGluLeu 322
DB 996 CCGAGCTGTGTCATTAAATCTGGCATCTGCGCGGAGAGAAATTCGCCCAACCTT 1055
QY 323 ValGluLeu 325
DB 1056 GAAGCGTTA 1064

RESULT 26
AP005089/c
LOCUS
DEFINITION Vibrio parahaemolyticus DNA, chromosome 2, complete sequence, 6/6.
ACCESSION AP005089 BA000032
VERSION AP005089.1 GI:28809850
KEYWORDS
SOURCE
ORGANISM
Vibrio parahaemolyticus
Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
Vibrionaceae; Vibrio.
REFERENCE
1 Nasu, H., Iida, T., Sugahara, T., Yamaichi, Y., Park, K.S., Yokoyama, K.,
Makino, K., Shingawa, H. and Honda, T.
A filamentous phage associated with recent pandemic Vibrio
parahaemolyticus O3:K6 strains
J. Clin. Microbiol. 38 (6), 2156-2161 (2000)
20295086
PUBMED
10834969
2
Makino, K., Oshima, K., Kurokawa, K., Yokoyama, K., Uda, T.,
Tagomori, K., Iijima, Y., Nakano, M., Nakano, M., Yamashita, A.,
Kubota, Y., Kimura, S., Yasunaga, T., Honda, T., Shinagawa, H.,
Hattori, M. and Iida, T.
Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
distinct from that of V. cholerae
Lancet 361 (9359), 743-749 (2003)
22508454
PUBMED
12620739
3 (bases 1 to 335862)
Hattori, M., Yamashita, A., Oshima, K. and Shiba, T.
Direct Submission
Submitted (09-APR-2002) Masahira Hattori, Kitasato Institute for
Life Sciences, Kitasato University; Kitasato 1-15-1, Sagamiara,
Kanagawa 228-8555, Japan
(E-mail: hattori@genome.ls.kitasato-u.ac.jp,
URL: http://genome.ls.kitasato-u.ac.jp,
Fax: 81-42-778-8193)
This work was done in collaboration with Kozo Makino, Ken Kurokawa,
Katsushi Yokoyama, Takayuki Uda, Kenichi Tagomori, Masatomo Najima,
Masayuki Nakano, Yoshino Kubota, Shigenobu Kimura, Teruo Yasunaga,
Takeishi Honda, Hideo Shinagawa, Tetsuya Iida (Osaka University),
Yoshio Iijima (Kobe Institute of Health), and supported by the
Research for the Future Program of the Japan Society for the
Promotion of Science. This clone was isolated from a patient
presenting with acute gastroenteritis.
Location/Qualifiers
1. 335862
/organism="Vibrio parahaemolyticus"
/mol_type="genomic DNA"
/strain="O3:K6"
FEATURES
source

```

```

/sub strain="RIMD 2210633"
/db_xref="taxon:670"
/chromosome="2"
/complement (55. .1971)
/gene="VPA1449"
/complement (55. .1971)
/gene="VPA1449"
/notes="similar to GB:AAF96089.1 (AE004358) percent
identity 71 in 638 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="methyl-accepting chemotaxis protein"
/protein_id="BAC62792.1"
/db_xref="GI:28809851"
/translation="MLKLSLSIKQKVLGITFAVLASTIIVGMAQRHARDVLSHRL
IDLEPAMLQIINTEIDREVMQQAOLATNEFVVEALKNTDHPQSFQVQQLN
NIKSOGLNDQAVANRKTAYVMNGFPLQNLNSQDAWPGFTSSGRETSVVFQAN
GEVKMTFYQDLNGISNGSLKSMDDVSLNLSFQIEDTGIFLTNKEGDIQIHRQOG
KNKTSIAQLQSGVANQLLNKSNFLNINVEFGKDNFIASLYPSNMWFIGVVPDEV
FALNATGQGMITIVVALVIFILMGVLLANSITTKPIRLIADFTDLGQEGDLSQRI
ERGNDEIAQLSGFNGFIEKIHATMKVEVSLTSGALSQAADSVSSTKATSYDNQSQR
DOTIIVAAINQMGATISIAANAATAADTANQASDNTOTGREVVMKAKEVLSRLADD
VETNIVVTQASTTKDIGSILGIVRDISEOTNLLALNALEAARAGEQGRGFAVVAD
EVRNLASRTADSPFEEIORMIQLQSDAQDAVSMEAGKAVTFEGVASTDEAVEVLVNI
SRITDISRNTQVATATEQSTVHTIINQNEEINAINEMTTATAEELAGASRLQIE
LSRLDMVGSFKL"
/complement (2136. .2288)
/gene="VPA1451"
/complement (2136. .2288)
/gene="VPA1451"
/notes="unknown"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC62793.1"
/db_xref="GI:28809852"
/translation="MLWTLDIHLIFVLSIGHQCSGKSCSLSKFTIFSILGQRKGLQAK
KNAESA"
2287. .3243
/gene="VPA1450"
2287. .3243
/gene="VPA1450"
/notes="similar to GB:AAF96088.1 (AE004357) percent
identity 79 in 318 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="MoxR-related protein"
/protein_id="BAC62794.1"
/db_xref="GI:28809853"
/translation="MHKSNFETLQHYLESQIIGQHDLVKQLLIALLADGHILVGGPPG
LAKTRAVKSLSDCVGEDFHRIOPTFDLLPADLTGDIIFRPSTGEFTFOSGPIFNSLIL
ADEINRAKQVQAMLEMAEKQVTAGRHVATLPELFVLMATNPIDEGTYPILPEAQ
LDRELLHLEVDPDAHELAIRINRGEAKGECIERPSLSQQDIPTARQEVLDIHW
EAIQYIVRLVMATRASEYSELDKWLAMGVSPRATIALDRCARAHAWLAGRDFVSP
EDVQAMYPVLRHLRLLLSYHAQAEVTSNQVIDKLLNLVGS"
3258. .4184
/gene="VPA1452"
3258. .4184
/gene="VPA1452"
/notes="similar to GB:AAF96087.1 (AE004357) percent
identity 50 in 309 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC62795.1"
/db_xref="GI:28809854"
/translation="MSANHPLPHANGVTLTLELLQYRTQSVRWLPPAQISQMSG
NHPISROKGRGMDPMEVRYQAGDDIRSIDMRVTARTKAHTKLPFAEDKEQAVILYLDL
CSSLHFGSKTVLKSQVLAHFASVLNLTAKKORIGAVIDDGHQCLCFRPSLSQKQGL

```

```

RILNAIVDTHNAQLTODPNHETRMPSOVLETLHTLTTPKGSSELILLSDFAHVBEREL
TQRLRLKQHNVSRAIQFYDPLERGETDFRQAKASDGRRSQWFGSTGORKKALETHF
TQHQESIQQOCHSMALFFNTLSSGWPLIQOLS"
4225. .4701
/gene="VPA1453"
4225. .4701
/gene="VPA1453"
/notes="similar to GB:AAF96086.1 (AE004357) percent
identity 51 in 151 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC62796.1"
/db_xref="GI:28809855"
/translation="MTTNQSNLEPLIILNAPDWFPLAWGMWMLAAIIVLFVLGVYL
YAKWRKRLALKTKLTLTQSVTPHTPSSALEILRQAALSYPREDIAPLTGMAYE
FIDSHTKTRFVMDKQOOWAALYQKSPQHQDLIDDCAFWITHALPPKKKANTRD"
4712. .5698
/gene="VPA1454"
4712. .5698
/gene="VPA1454"
/notes="similar to GB:AAF96085.1 (AE004357) percent
identity 65 in 317 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC62797.1"
/db_xref="GI:28809856"
/translation="MTDLSQMLNIEFVMWALFLPLPLIYKLLPEETROAEIKLAY
LPESKNSRKPQWLQKSLAVGVWTLIVACAPVWFGPEVFPKYRDLMLVVDLSGS
WQKEDMNLDEYIDRLSAVKKVLSDFAVRKGRDLGVLFGDHAYLQTLPTADRTVI
QOIQTIVGLVQRTAIGDIGLGTFTFVSDAPQRMVILSGNTAGVLDPIEAAE
IAKKNYATITVGVGAGEMVMKDFPMTRKVDTAADLDEQTLTKIAEMTGGQYFRADA
EQLEKIYDTINKLEPVSDTOTWRPQSEWFPYPLSALVLSLFLLRKKG"
5691. .7547
/gene="VPA1455"
5691. .7547
/gene="VPA1455"
/notes="similar to GB:AAF96084.1 (AE004357) percent
identity 43 in 653 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC62798.1"
/db_xref="GI:28809857"
/translation="MASFTPLNPLYLLGLITIPFVLLINQKFRAKKSALIAPIHLAKML
GOTTQSKPYFTINGLAWTIACLAGPSWQNTSPSFELSONRMLVLDMSRSMYASDI
KPNLAQTRYKALDILLPKWKEGATGLIIVYAGDAYLSPLTTDSTLAGIINLSPELM
PFQSNLPAIAIELSQFSQAGANQGDIVVLADDLDSSELSLDLVKGNIRVSVLA
IGTAGAPIALPDGSLKTKTQGSTVAKTKLNLQTLANKTGMGFVPIQHNNRVDVSI
AAVTNNTGNLAAKQNOEVKTDLSRLNGFWLLPLLLPTVPLFRGVIVVFAALLPW
TWTDAQANPLNADQAAANLYKQGEYEQAKNLTDFDSWQGAASYQTGDYBAALKAFS
NDPSQTRYNLANAQNGLEADKQYKLLLENPDFEAAKNLSVVEEKLKQOQQQ
QOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQO
QOQQOQNRTKQEQSQNQDQENKEKEQSEHKAQKQKAEVEKQEQKQGPMPQ
AOAQEKSHODDPFRLEAVESARDPSFLIRAQMLQAOQKQRPQSQKEM"
7555. .9198
/gene="VPA1456"
7555. .9198
/gene="VPA1456"
/notes="similar to GB:AAF96082.1 (AE004357) percent
identity 44 in 200 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC62799.1"
/db_xref="GI:28809858"
/translation="MKKGVLVFLVASLLSSFVSVAQSLQASVNNKTEVAKNEVINUR
IMADSEVASDAIDFSVLEKDFLQGPYRGKSSNNNGRKRYQRTIEWSISIAPMKEGIIT

```



US-10-724-972A-6352 (1-335) x BX950851 11 (1-110000)

QY	8	LeuLysIleLeuSerValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCys---	26
Db	105548	ATTAGAGTATTCCTTATGTTGCTCTACTCTCGCTTGCATGATTACAAATGCTTTGTCGTG	105489
QY	27	-----GlyAsnAenSerSerAenSerSerLysGluSerSerLysAspGlyVal	43
Db	105488	TTCCGACTGGGTCCG-----GCCAATGCAGTA	105462
QY	44	GluIleLysHiGluGluGlyThrThrLysValProLysHiAProLysArgValValVal	63
Db	105461	ACGGTGCAGGATGAACAGGGTTCCTTACGCTCGACACGCCCTCAGCGCATTTGGTG	105402
QY	64	LeuGluTySerPheValAspAlaLeuValAlaLeuAspValLysProValGlyLeuAla	83
Db	105401	CTGGAACTCTCTTTTGGCCGATGCATGGCCGCGATAGACGTGAGCCCATCGCATCGCC	105342
QY	84	AspAspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyThr	103
Db	105341	GATGACACGATCCAGACCGGATCTGGCAGAAAGTCCGTGAACACCTAAACCGTGGCAC	105282
QY	104	SerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeu	123
Db	105281	TCAGTGGGAACCGTGCAGCGCGAGCCTCGAAGCCATCAGCGCTGAAGCCTGATTTG	105222
QY	124	IleIleAlaAspAsnAsnArgHisLysGlyIleTyLysAspLeuAsnLysIleAlaPro	143
Db	105221	ATTATCGCCGACAGCAGTCGCCACAGCGGTATTTATACGGCACGTGAAGCTATCGTCCG	105162
QY	144	ThrIleGluLeuLysSerPheAspGlyAspTyAsnGluAsnIleAspAlaPheLysThr	163
Db	105161	GTACTGCTGCTGAATCAGCAACCACTATGATGAACACCTGCTCATTCGGCGAATC	105102
QY	164	IleSerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHiAspLys	183
Db	105101	ATCGCAAGAGTATTAGGGAAACACAGTCGATGCAAGCTCGTCTTGTCTAAGCATCGCGC	105042
QY	184	LysIleGluGluTyLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAla	203
Db	105041	ACCATGAAGCTACGCAGACCAG-----CTCCAAAGCACGAGCGCTTCGCGC	104988
QY	204	ValAlaAlaLysSerSerGlyLeuLeuAlaHisProSerAsnSerTyValGlyGlnPheLeu	223
Db	104987	ACCTCGCGCAACAGCAGTTTAACTGTGATTCCAGCGAGGCTACACCGGACGGTGTG	104928
QY	224	SerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTy	243
Db	104927	GCTGCTCTGGGTCTG-----AGCGTGCAGAAACCATCAAT-----	104892
QY	244	LeuLysGlyProTyLeuGlnMetAsnThrLysLeuSerGlnValAsnProGluArg	263
Db	104891	---AATGCCCGATGGCTCCATTCATTTGGACAGCTGCTGGCGATTAATCCACAGTG	104835
QY	264	MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGluLys	283
Db	104834	CTGATTGTGACGCCTATCTCG-----GAAGAAAGTATTGTGAACCGTGGCAGCAG	104784
QY	284	AspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg	303
Db	104783	GATACGCTCTGGAAACATGCTGGAAAGCACAGCAAAACAGCAGATTGCGCGCTGCAGCT	104724
QY	304	AspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuVal	323
Db	104723	AACGCTGGGGCGGATCGCGGCATCTTTGACGCGACGCGTTCGCGCAGCATCGGTG	104664
QY	324	GluLeuSerLys	327
Db	104663	AAGATTTTTTAA	104652

RESULT 28  
BX950851\_12/c  
WPCOMMENT

Sequence split into 51 fragments LOCUS BX950851 Accession BX950851

Fragment Name	Begin	End
BX950851_00	1	110000
BX950851_01	100001	210000
BX950851_02	200001	310000
BX950851_03	300001	410000
BX950851_04	400001	510000
BX950851_05	500001	610000
BX950851_06	600001	710000
BX950851_07	700001	810000
BX950851_08	800001	910000
BX950851_09	900001	1010000
BX950851_10	1000001	1110000
BX950851_11	1100001	1210000
BX950851_12	1200001	1310000
BX950851_13	1300001	1410000
BX950851_14	1400001	1510000
BX950851_15	1500001	1610000
BX950851_16	1600001	1710000
BX950851_17	1700001	1810000
BX950851_18	1800001	1910000
BX950851_19	1900001	2010000
BX950851_20	2000001	2110000
BX950851_21	2100001	2210000
BX950851_22	2200001	2310000
BX950851_23	2300001	2410000
BX950851_24	2400001	2510000
BX950851_25	2500001	2610000
BX950851_26	2600001	2710000
BX950851_27	2700001	2810000
BX950851_28	2800001	2910000
BX950851_29	2900001	3010000
BX950851_30	3000001	3110000
BX950851_31	3100001	3210000
BX950851_32	3200001	3310000
BX950851_33	3300001	3410000
BX950851_34	3400001	3510000
BX950851_35	3500001	3610000
BX950851_36	3600001	3710000
BX950851_37	3700001	3810000
BX950851_38	3800001	3910000
BX950851_39	3900001	4010000
BX950851_40	4000001	4110000
BX950851_41	4100001	4210000
BX950851_42	4200001	4310000
BX950851_43	4300001	4410000
BX950851_44	4400001	4510000
BX950851_45	4500001	4610000
BX950851_46	4600001	4710000
BX950851_47	4700001	4810000
BX950851_48	4800001	4910000
BX950851_49	4900001	5010000
BX950851_50	5000001	5064019

Alignment Scores:

Pred. No.:	1.16E-24	Length:	110000
Score:	477.50	Matches:	110
Percent Similarity:	54.01%	Conservative:	65
Best Local Similarity:	33.95%	Mismatches:	120
Query Match:	28.23%	Indels:	29
DB:	1	Gaps:	6

US-10-724-972A-6352 (1-335) x BX950851 12 (1-110000)

Qy	8	LeuLysIleLeuSerValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCys---	26
Db	5548	ATTAGGAGTTACGCTATTTGCTCTACTTCGGGTGCACATGATTAACAATGCTTTGTCTG	5489
Qy	27	-----GlyAsnAsnSerSerAsnSerSerLysGluSerLysAspGlyVal	43
Db	5488	TTGCGACTGGGTGCG-----GCCAATGCAGTA	5462



Db 7829 CAGGATACCTGTCGCAGATGATCAGCAGCAGCAGCAGCATCAGATGATGACGCGGTAGAC 7770  
Qy 303 ArgAspLeuTrpAlaAaGserArgGlyLeuIleSerSerGluGluMetAlaLysGluLeu 322  
Db 7769 AGCAACACCTGGCTGCTGTCGCGGGATATTTCGCCGCTGAACCATGTCGACGCGATGCG 7710  
Qy 323 ValGluLeu 325  
Db 7709 GTAATAATC 7701

RESULT 30  
BX571873  
LOCUS  
DEFINITION  
Photorhabdus luminescens subsp. laumondii T701 complete genome;  
segment 15/17.  
ACCESSION  
BX571873 BX470251  
VERSION  
BX571873.1 GI:36787441  
KEYWORDS  
complete genome.  
SOURCE  
Photorhabdus luminescens subsp. laumondii T701  
Photorhabdus luminescens subsp. laumondii T701  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Photorhabdus.

REFERENCE  
AUTHORS  
Duchaud, E., Rusniok, C., Frangeul, L., Buchrieser, C., Taourit, S.,  
Bocs, S., Boursaux-Eude, C., Chandler, M., Dassa, E., Deroose, R.,  
Dorzelle, S., Freyssinet, G., Gaudriault, S., Givaudan, A., Glaser, P.,  
Medigue, C., Lanois, A., Powell, K., Siguier, P., Wingate, V.,  
Zouine, M., Boemare, N., Danchin, A. and Kunst, F.  
Complete genome sequence of the entomopathogenic bacterium  
Photorhabdus luminescens  
Nat. Biotechnol. 11 (1) (2003) In press

JOURNAL  
REFERENCE  
AUTHORS  
Duchaud, E., Frangeul, L., Rusniok, C. and Kunst, F.  
Direct Submission  
Submitted (23-APR-2003) L. Frangeul, Institut Pasteur, Genopole, 25  
rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE. E-mail:  
lfrangeu@pasteur.fr, fkunst@pasteur.fr  
LOCATION/Qualifiers  
1. 349107  
/organism="Photorhabdus luminescens subsp. laumondii T701"  
/mol\_type="genomic DNA"  
/strain="T701"  
/db\_xref="taxon:243265"  
complement (80. .3225)  
/gene="tccc1"  
/locus tag="plu4167"  
complement (80. .3211)  
/gene="tccc1"  
/locus tag="plu4167"  
/function="Unknown, probable insecticidal toxin"  
/note="Highly similar to insecticidal toxin complex  
protein TccC of W14."  
/codon\_start=1  
/transl\_table=11  
/product="Insecticidal toxin complex protein Tccc1"  
/protein\_id="CAE16539.1"  
/db\_xref="GI:36787442"

translation="MSTPDTALYTQPTVSLVDRGLSIRDIGPHRVVIGGIDITRVT  
RHOVDARGYLNSIDPLRYDAKQADNSVKNFVQYDLGHALTESEVDAGRTVALND  
IEGRPVMTNATGVRQRRYEGNTLPGRLSSVQVDFQETSQVTERFIAENTTAEK  
EYNSGTCIRHYATGVTGRLMSQSLAGAILSQSHQWLAEQANSGDETVMQMLA  
SEVYTTSTTNAIGALLTQDAKNIQRLAYDIAGQLKSWLTVKGQEOVIVKLSW  
SAAGKLEHNGGVITEYSYEPETQRLMGTTRRAESQSEARVLODLRYEYDPVGN  
VISIHNDAEATFRWRNOKVEPENRYVYDSLQVLSATGREMANIQGSNQLPSPAMPV  
PADNSTYNTQRTSYDRGNLVQIRHSSPATQNSYTMVSRSNRNTVLSTLTTDP  
TQVGLFDAGHQKTLIPGNLNDWNRGELQRTVPVSRNSSESWRYSSDGVRLIK  
VSEQTQNSQVQRTVLPGLRLITGVADTKTQNLQVITVGEAGRAQVRVLHWSGK  
PADIDNNQVRSYDNLGSSQIELDSEQLDSOEYYPVGLNVAANKNOTEASXFI  
RYSCKERDAGLYYGYRYPQWRLSDPAGTVPGLNVRNRPITLTDHGL  
APSPNRNNTFWPASFPLFRKPDGMSMRGOKIGRALAGGLAIGSLAAIAGTAGA  
ALPVLGVAAGAGIGALMGVNSLLEKGCALLARLVQGRKSTLVQSAAGAAGASSA  
AAYGARAGQGVGASAGAVTGAVGSINNADRGIGGAGAGAGATGTTMLGTSSTLT  
HEVAAGAAGGAGGMITQGSTAGIHAGITGYTGYSGWIGFGLDVASNPAGHLANYAG

translation="MSTPDTALYTQPTVSLVDRGLSIRDIGPHRVVIGGIDITRVT  
RHOVDARGYLNSIDPLRYDAKQADNSVKNFVQYDLGHALTESEVDAGRTVALND  
IEGRPVMTNATGVRQRRYEGNTLPGRLSSVQVDFQETSQVTERFIAENTTAEK  
EYNSGTCIRHYATGVTGRLMSQSLAGAILSQSHQWLAEQANSGDETVMQMLA  
SEVYTTSTTNAIGALLTQDAKNIQRLAYDIAGQLKSWLTVKGQEOVIVKLSW  
SAAGKLEHNGGVITEYSYEPETQRLMGTTRRAESQSEARVLODLRYEYDPVGN  
VISIHNDAEATFRWRNOKVEPENRYVYDSLQVLSATGREMANIQGSNQLPSPAMPV  
PADNSTYNTQRTSYDRGNLVQIRHSSPATQNSYTMVSRSNRNTVLSTLTTDP  
TQVGLFDAGHQKTLIPGNLNDWNRGELQRTVPVSRNSSESWRYSSDGVRLIK  
VSEQTQNSQVQRTVLPGLRLITGVADTKTQNLQVITVGEAGRAQVRVLHWSGK  
PADIDNNQVRSYDNLGSSQIELDSEQLDSOEYYPVGLNVAANKNOTEASXFI  
RYSCKERDAGLYYGYRYPQWRLSDPAGTVPGLNVRNRPITLTDHGL  
APSPNRNNTFWPASFPLFRKPDGMSMRGOKIGRALAGGLAIGSLAAIAGTAGA  
ALPVLGVAAGAGIGALMGVNSLLEKGCALLARLVQGRKSTLVQSAAGAAGASSA  
AAYGARAGQGVGASAGAVTGAVGSINNADRGIGGAGAGAGATGTTMLGTSSTLT  
HEVAAGAAGGAGGMITQGSTAGIHAGITGYTGYSGWIGFGLDVASNPAGHLANYAG

YAAAGLGAEMAINRVIGGGPLGRLLGRAVSPYAAGLARQLVHFSIARVPFPIFVLGG  
LAGGIGTGLHRVWGRDSWSRGLSAAAGSDIHLAGMIGNQIRGRVLTTTGIANAIDYG  
TSAGAAARRVISL"  
complement (3220. .3225)  
/gene="tccc1"  
/locus tag="plu4167"  
complement (3345. .8051)  
/gene="tccc1"  
/locus tag="plu4168"  
complement (3345. .8039)  
/gene="tccc1"  
/locus tag="plu4168"  
/function="Unknown, probable insecticidal toxin"  
/note="Highly similar to insecticidal toxin complex  
protein tccB of W14 Putative transmembrane protein."  
/codon\_start=1  
/transl\_table=11  
/product="Insecticidal toxin complex protein TccB1"  
/protein\_id="CAE16540.1"  
/db\_xref="GI:36787443"  
translation="MLSTMEKQLNESQDALVTGYMNFVASTLKGVDGPVTVEDLYE  
YLLIDPEVADEVETSRVQAIAISIQYMTLVNGSEPGRAQMELSANERWDSOVA  
VNSLVLSGYINQDKFDOAIYFIFGRTTTPKYRYWRQDLSDRQVODAVLYLNERFEA  
NDWQETLPLSGDVLVLEHTVRPFVFNDRLYVAVVERDPAQVQADGKNIKTHAYSIK  
FGYKYDDTPAPNTTLMTQADSESQTSRSLIDESKTLTKRVLNLLATDFFSIDP  
TEETDSNPYGRMLMGVFNRTFTGPHQNPVAVLYLDCDSAFNHHILRLPLSDFFST  
YRDETNDQCQFQAVYDKYVITKISGPTDPENTGWGVDDLLKLTGTGAYVTFH  
QDGLTHIQITTTNGDFVNRHIFGYNDIVVDSKSGYFTWSGNEGFVLDYVHDGYTFH  
NAAINYPAGYGGSGVPGNCTWALEQVNEGWAIPLDDAHTTITVQGSIAWEGETPA  
SYLPIPDGTMLLDMSDKINFAISLKLKLESVFTSPDWPLTTIKFSLKIDNRKFKQE  
INAEADGNLFPKHYSQTFQALAGATYPTTTLSEADFPSTDPKXNLQVCLNVWDH  
YDRSGRGASVSWSNWVYVVELQDSKAPDAIPRLISRYDSKRGVLQVDFWTSLLP  
AKRLNTTFTVTLIEKANGLDLSLDYTLQADPSLEADLVTDKSEPMDFDGSNGLYF  
WELFPLPLVATRFANEQQSPQAQSLHYIFDPAMKPNHNAWVNRVPLVEHSD  
LSRHLGDPIDPDTQAYAPVYQKAVFIAYVSNLIAQGDWYRQLTRDGLTOARVYIN  
LAELGPRPDSVLSSTWTPQTLAAGQKTLRDFEHQLANSVTLPALPENNISH  
LKLADNGYFNEPLVLMSLHWDTLDRNLNHNLTVDGKPLSLPLYAAPADPVALLA  
QRAQGTUANGSVGAILTVPRFSAMLPRAYSVGTTLTSFQNLISLLEPESERAGQE  
BLAQQLDMSYAITLQOQALDGLADRLALLASQATQQRHHDHYTLQNNISAE  
QLVMDTQTSQSLISSSTGVOTAGALKVIPNIFGLADGSRYEGVTEAIALGLMAAG  
QATSVVAERLATTENRRREWEQIOYQAOSEVDALOKOLDALAVREKAAQTSIQQA  
KAQVQIRMTLYLTTRTQATLQWLSQLSALYQAYDVAUSLCSAQCWQVELG  
DYATFTQGTWNBHYRGLQVGETLQNLQMEAAVLRHERRVNLIVTSLSLGD  
DQFKLTKEGKVDPLSEKLPDNDYPGHYLRQITVSVTLPTLIGPQNVKATLTQTS  
STLLTADINGKRLNDTPKGEQATHIVTNLRASQVQVALSSGINSAGSFEURLDER  
YLSFEGTGAVSKWTLTTPRSVDEHIDDKTLKADQMQAALLANMDDVLQVHYTACDGG  
ASFANQVKTLF"  
complement (8046. .8051)  
/gene="tccc1"  
/locus tag="plu4168"  
complement (8134. .11034)  
/gene="tccc1"  
/locus tag="plu4169"  
complement (8134. .11034)  
/gene="tccc1"  
/locus tag="plu4169"  
/function="Unknown, probable insecticidal toxin"  
/note="Highly similar to insecticidal toxin complex  
protein TccA of W14."  
/codon\_start=1  
/transl\_table=11  
/product="Insecticidal toxin complex protein TccA1"  
/protein\_id="CAE16541.1"  
/db\_xref="GI:36787444"  
translation="MNQLASPLISRTSEIHNLPKLTDLGYTSVDFVVRMPREPPIRE  
HRADLGSASKMVDLAVGYAHVHLHFRNLSSEAVQFSURSPFSVSGPDYANFLDA  
NTGWCKAPSGSPENDAPVAYLTHIYQLADQKSGASAIMNTLVERRPDIGALLIN  
DAKINEVILPOLVNEILSKAIQKLSLTDLEAVNAKLSTTRPNNLPYHYHQIOIOT  
AQVSLTLDITLTPQTLDPQNFATAKGLKSDATAGALTRLIQMASOLSPEQOKIM  
VETVQGVNFCYQGLSLLTANFSNDLTDRSLTVPQVELMLCSTVSGSVYKSDN  
VSSGDTTATPFAYGARFIHAGKQATLSRSRGAHFPALVNLNLDKLDORINTKRL  
QKWLNPYEDIDLIVTSAMDDETNTALSMDNDTLRLMLGVFKHYQAKYGTAKQFAGH

DB:	1	Gaps:	3
US-10-724-972A-6352 (1-335) x BX571873 (1-349107)			
Qy	43	ValGluLeuLeuHisGluGluGlyThrThrLysValProLysHisProLysArgValVal 62	
Db	331083	GTCCGGTACAGGATGAGCAGGCGAGTTTACCGTCAATACCGTTCCCGACGGGTAGTC 331142	
Qy	63	ValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyLeu 82	
Db	331143	GTGCTGGAATGTGATTGGCGATGCAATGGCGCCCAATATACGTCCCGTAGGTATT 331202	
Qy	83	AlaAspAspAsnLysLysAsnArgLileLysProLeuArgAspLysAlleGlyLysTyr 102	
Db	331203	GCCGATGATAACGATCCGACGCGCACTTCACCGCATGTCGCCAGCGTATTAAAGCCCTGG 331262	
Qy	103	ThrSerValGlyThrArgLysGlnProAsnLeuGluGluLileSerLysLeuLysProAsp 122	
Db	331263	CAATCAACCGGCACTCCGCGCCCAACCAAGCCCTGAAGCCATCAGCGCGCTGAACCTGAT 331322	
Qy	123	LeuLileAlaAspAsnArgHisLysGlyLileTyrLysAspLeuAsnLysLileAla 142	
Db	331323	CTCATATTGCAGACAGTCAACGCCACCGCGGGATTTATCGGCCACTCAAGGGCATCGCT 331382	
Qy	143	ProThrLileGluLeuLysSerPheAspGlyAspTyrAsnGluAsnLileAspAlaPheLys 162	
Db	331383	CCTGCTGCTGCTAAAAATCCGCAACCAAAACCTACGAGAGAACCTGCAATCTCGCGCG 331442	
Qy	163	ThrLileSerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAsp 182	
Db	331443	ATCATCGGCAAGTCTGCGGAAAGAGAGTGCAGTGCAGAAACCGCTGCAGCAGCACCGT 331502	
Qy	183	LysLysLileGluGluTyrLysLysGluLileThrMetAspLysAsnGlnLysValLeuPro 202	
Db	331503	GAACGATGAAAGGCTATGCCAGCCAG-----TTGCCACAAAGGCGTCACTGTGTTT 331556	
Qy	203	AlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe 222	
Db	331557	GGCACCTCGCGCGAAACGAATTTAACTGCACTCCAGGATACCTATACCGCAGCGTC 331616	
Qy	223	LeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLys 242	
Db	331617	TTAACAGCATTAGGCGTAAAGTTCCCGCATCAGTTAAT-----331655	
Qy	243	TyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGlu 262	
Db	331656	-----CATGCGGCAATGGTTTCTCTCAACCTGGAAACAGTTACTGCATCACTCTGAC 331705	
Qy	263	ArgMetPheLileMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGlu 282	
Db	331710	TGGCTTATCGTGGCCCACTATCGT-----CAGAAAGCATCGTGAACCGCTGGCAG 331760	
Qy	283	LysAspProValTyrLysLysLysLeuAsnAlaValLysAsnGlnArgValAspLileLeuAsp 302	
Db	331761	CAGGATATCTGTGGCATGATGATGACAGCAGCAGCAGCATCAGATGATAGCAGCGGTAGAC 331822	
Qy	303	ArgAspLeuTyrAlaArgSerArgGlyLeuLileSerSerGluGluMetAlaLysGluLeu 322	
Db	331821	ACCAACACCTGGCTCGTATGCGCGGATATTTCGCGTGAACGCATTTGCGAGCGATGCG 331888	
Qy	323	ValGluLeu 325	
Db	331881	GTAATAATC 331889	
RESULT 31			
AB354237			
LOCUS	AB354237	668 bp	DNA linear PAT 17-AUG-2003

Alignment Scores:		
Pred. No.:	5,99a-24	Length:
Score:	476.00	Matches:
Percent Similarity:	57.24%	Conservative:
Best Local Similarity:	36.04%	Mismatches:
Query Match:	28.17%	Indels:
		349107
		102
		60
		107
		14

AR354237	AR354237	668 bp	DNA	linear	PAT 17-AUG-2003
LOCUS	Sequence 355 from patent US 6593114.				
DEFINITION	AR354237				
ACCESSION	AR354237.1				
VERSION	GI:33760321				
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unknown.				



Unclassified.  
1 (bases 1 to 668)  
Kunsch, C.A., Choi, G.H., Barash, S., Dillon, P.J., Fannon, M.R. and  
Rosen, C.A.  
TITLE  
Staphylococcus aureus polynucleotides and sequences  
JOURNAL  
Patent: US 6593114-A 355 15-JUL-2003;  
FEATURES  
Location/Qualifiers  
1. .668  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Alignment Scores:  
Pred. No.: 3.9e-27 Length: 668  
Score: 475.00 Matches: 91  
Percent Similarity: 86.89% Conservative: 15  
Best Local Similarity: 74.59% Mismatches: 16  
Query Match: 28.11% Indels: 0  
DB: 6 Gaps: 0

US-10-724-972A-6352 (1-335) x AR354237 (1-668)  
QY 210 LeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLys 229  
Db 2 TTATTAGCACATCCAACTATTCTATGTTGGACAATTTTAAACGAAGTAGGATTTAAA 61  
QY 230 GluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeu 249  
Db 62 AATGCATTAAAGTACGATGTACAAAGGTTTAAGTAAATATTGAAAGACCTTACTTA 121  
QY 250 GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn 269  
Db 122 CAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAGCGTATGATTCATTATGACAGAT 181  
QY 270 LysAlaSerSerAsnGluProSerLeuLysGluLeuLysAspProValTrpLysLys 289  
Db 182 CATGCTAAAAAGATTCTCTGAAATTCAGAAGATTACAAGAAGATGCAACATGGAAAAAG 241  
QY 290 LeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSer 309  
Db 242 TTGAATGCATTAAAAAATAATCGCTGGATATTGTTGACCGGATGTTTGGGCAAGATCT 301  
QY 310 ArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAsp 329  
Db 302 COTGCTTTAAATTTCTTCTGAAGAATGCTTAAGAAGACTTGTGTAATTATCAAAAAAGAA 361  
QY 330 SerLys 331  
Db 362 CAAAAG 367

RESULT 32  
LOCUS AR535793 668 bp DNA linear PAT 08-OCT-2004  
DEFINITION Sequence 355 from patent US 6737248.  
ACCESSION AR535793  
VERSION AR535793.1 GI:53927010  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 668)  
Kunsch, C.A., Choi, G.A., Barash, S.C., Dillon, P.J., Fannon, M.R. and  
Rosen, C.A.  
TITLE  
Staphylococcus aureus polynucleotides and sequences  
JOURNAL  
Patent: US 6737248-A 355 18-MAY-2004;  
FEATURES  
Location/Qualifiers  
1. .668  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Alignment Scores:  
Pred. No.: 3.9e-27 Length: 668

Score: 475.00 Matches: 91  
Percent Similarity: 86.89% Conservative: 15  
Best Local Similarity: 74.59% Mismatches: 16  
Query Match: 28.11% Indels: 0  
DB: 6 Gaps: 0

US-10-724-972A-6352 (1-335) x AR535793 (1-668)  
QY 210 LeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLys 229  
Db 2 TTATTAGCACATCCAACTATTCTATGTTGGACAATTTTAAACGAAGTAGGATTTAAA 61  
QY 230 GluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeu 249  
Db 62 AATGCATTAAAGTACGATGTACAAAGGTTTAAGTAAATATTGAAAGACCTTACTTA 121  
QY 250 GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn 269  
Db 122 CAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAGCGTATGATTCATTATGACAGAT 181  
QY 270 LysAlaSerSerAsnGluProSerLeuLysGluLeuLysAspProValTrpLysLys 289  
Db 182 CATGCTAAAAAGATTCTCTGAAATTCAGAAGATTACAAGAAGATGCAACATGGAAAAAG 241  
QY 290 LeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSer 309  
Db 242 TTGAATGCATTAAAAAATAATCGCTGGATATTGTTGACCGGATGTTTGGGCAAGATCT 301  
QY 310 ArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAsp 329  
Db 302 COTGCTTTAAATTTCTTCTGAAGAATGCTTAAGAAGACTTGTGTAATTATCAAAAAAGAA 361  
QY 330 SerLys 331  
Db 362 CAAAAG 367

RESULT 33  
LOCUS ECOFECBCDE 4842 bp DNA linear BCT 26-APR-1993  
DEFINITION K-12 fecA gene, 3'end; fecB,C,D, and E genes, complete cds's.  
ACCESSION M26397  
VERSION M26397.1 GI:145923  
KEYWORDS  
SOURCE Escherichia coli  
ORGANISM Escherichia coli  
REFERENCE  
1 (bases 1 to 4842)  
Staudenmaier, H., Van Hove, B., Yaraghi, Z. and Braun, V.  
Nucleotide sequences of the fecBCDE genes and locations of the  
proteins suggest a periplasmic-binding-protein-dependent transport  
mechanism for iron(III) dicitrate in Escherichia coli  
J. Bacteriol. 171 (5), 2626-2633 (1989)  
89213950  
2651410  
COMMENT Original source text: Escherichia coli (strain K-12) DNA.  
FEATURES  
Location/Qualifiers  
1. .4842  
/organism="Escherichia coli"  
/mol\_type="genomic DNA"  
/strain="K-12"  
/db\_xref="taxon:562"  
1. .136  
/gene="fecA"  
1. .136  
/gene="fecA"  
/codon\_start=2  
/transl\_table=11  
/protein\_id="AAA23761.1"  
/db\_xref="GI:145924"  
/translation="DLNLAFGVKNI PDQYFIRSYDDNNKGIYAGQPTLYMQSLKP"  
181. .1083  
gene



AUTHORS Luck,S.N., Turner,S.A. and Rajakumar,K.  
 TITLE Direct Submission  
 JOURNAL Submitted (06-DEC-2000) Microbiology, Monash University, VIC 3800, Australia  
 REFERENCE 3. (bases 1 to 66714)  
 AUTHORS Luck,S.N., Turner,S.A. and Rajakumar,K.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-MAY-2002) Microbiology, Monash University, VIC 3800, Australia  
 REMARK Amino acid sequence updated by submitter  
 REFERENCE 4. (bases 1 to 66714)  
 AUTHORS Luck,S.N., Turner,S.A. and Rajakumar,K.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-JUN-2002) Microbiology, Monash University, VIC 3800, Australia  
 REMARK Sequence update by submitter  
 COMMENT On Jun 18, 2002 this sequence version replaced gi:15808696.  
 FEATURES  
 source location/Qualifiers  
 1..66714  
 /organism="Shigella flexneri 2a"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:42897"  
 /note="corresponds to Escherichia coli min 23"  
 misc\_feature 458..66714  
 /note="SRL pathogenicity island"  
 gene 619..1836  
 /gene="int"  
 CDS 619..1836  
 /gene="int"  
 /note="member of the P4 bacteriophage integrase family"  
 /codon\_start=1  
 /transl\_table=11  
 /product="integrase"  
 /protein\_id="AAL08426.1"  
 /db\_xref="GI:15808697"  
 /translation="MAVLDTKARHKPPDKPLPHGGITGLTLHPSSVKRGKWKVPFA  
 VSPVTKRRNAGLTGYDEVSIAGEAARTAIRMRQLAAGDDPLBIKKAEEKVVIPTFA  
 DAARRVHAELSGPWENPKHVMQSLTLENVAFPQLGAKTLDSTIADVAETLAPWLIT  
 LSTASRVKORIHVMQWMAWGFCVANPVDVHLLPQOTRGDRDEHPAMPWRQQLP  
 FVATSVYSDPYNVTRALLMWLITATRSGEARGMWAEIDFHKRVMTTIPAERMKARL  
 OHVPLSRQIYILENIRIGLDEALPVPSPKQIILSDMVLTSFLRKKKAVSDITPGRVA  
 TANGFRSTFRDWCSEOGYSRDLAERALAHTLKNKVEAAVHRTDLELRQVRPMMQAWAY  
 VMSQLAKNTKEFA"  
 CDS complement(2006..3535)  
 /note="orf2; similar to yfj1 on the cryptic P4 prophage  
 CP4-57 of Escherichia coli K-12"  
 /codon\_start=1  
 /evidence=not experimental  
 /transl\_table=11  
 /product="unknown"  
 /protein\_id="AAL08427.1"  
 /db\_xref="GI:15808698"  
 /translation="MMTKLSLPELDTTKLPDRCQALLDMHEETGIRREILLSSVLT  
 VMAASVQDTHREVELSGQRTSLQIFMCLSSASGSGKTSACAKLIAPVHETEELHQA  
 IDDKNYNDPENMMWTTDKTILERRYKKEITERPENAAARAALKETCIANKPPVPVQV  
 LLVNDATPEGIALKLSQSPSLILLDEGGTILDRPERKSAIYNTLWSCQPVNTERAS  
 RQGFIRLVNLTILITQPIVFNKFTLTGDIQRNGFLARLLFCPEGAIKITETPEH  
 TATPVVTOQACSGCFKSGIQSDLSRASRRRAKSKERI CMTLSNAASRAILYFIE  
 ENMDTVRQPHMGTFEDIIVRKEQVVRIAALLELEKDPDSTVITLESNTSAISAIYFIE  
 YFKFLYKLESLREISFPAEKLQWLOKRIITTAGYIFQKSYILOVAPVALRKCKVLDL  
 ALDILAEQRKIRIDNNLVVIGNTITPSELAKKLNIPAFDAGVFCIDHQNTILKYHNR  
 L"  
 CDS complement(3507..3812)  
 /note="orf3; similar to AlpA on the cryptic P4 prophage  
 CP4-57 of Escherichia coli K-12"  
 /codon\_start=1  
 /transl\_table=11  
 /product="putative transcriptional regulator"  
 /protein\_id="AAL08428.1"  
 /db\_xref="GI:15808699"  
 /translation="MHMMNMNSGNTQGYTPSRIILRMABLTRLLGISRSTIYVKINPA  
 SKYHDPTFPKPIRLGASGVGRASAIDEWMLHTAPAWSEPNKLEKLDNDDDKTITP"  
 3868..3873  
 CDS

/note="orf9; similar to DcuA in Escherichia coli K-12"  
/codon\_start=1  
/transl\_table=11  
/product="putative anaerobic decarboxylate transporter"  
/protein\_id="AAL08434.1"  
/db\_xref="GI:15808705"  
/translation="MWELIIVLGAIFRIGQGGIGLGGGLGAILTLTGFLPMG  
SPVDVILLIMTVVAASALQAAGMDYLVRVAGNPMRNPXYINIIPIITWSMTIM  
AGTGFVFTSLPIAEVAKESGIRPSTLAGVSQVALSGSPISAAAAMLTINEV  
NGVSFTQMSVCLPTSFAAAMVAAPFASQGGELQDDVYLERLQGLVKYENNSSI  
KPCATLUSGLFLATVAIVLAAFPQIRPGFDIGKPMETRDIIICMLSAACLMVLC  
KYATDAIILTSFRAGMSLAVILGIVTGTTFIDAHLEIKDIAIGDILQAPVPMIAL  
VIFGTCALYSQAGTPLIPLAVALDVPTMAILASVAVTGVFVLPVTPVTSLAAMEF  
DTGTTRVKYVNLNHPFMLPGUGGIITAGVALGVIAPMIL"  
/complement(<14553..>15322)  
/notes="remnant of 15600-like element; central region  
deleted"  
repeat\_region  
complement(15323..16090)  
/note="ISI element"  
/complement(16171..16962)  
gene  
/genes="aadA1"  
CDS  
/genes="aadA1"  
/note="confers streptomycin and spectinomycin resistance;  
similar to AadA1 from Tn21 on plasmid R100"  
/codon\_start=1  
/transl\_table=11  
/product="streptomycin adenyltransferase AadA1"  
/protein\_id="AAL08435.1"  
/db\_xref="GI:15808706"  
/translation="MRVAVIAEYSTQSEVVGVIETRLHLLAVHLVGSVDGLKP  
HSDIDLVTVRDLTERRALLNDLSTASPGESILRAVEVTVVHDDILPWRYP  
AKRLOFGWKXNDILAGIFEPATIDIDLAILLKAREHVALVGPAAELFDFVPQ  
DLPEXNETLTMNSPDWQXRVNVTXIXWXTVXIAPKDVXADWAMERLPA  
QYQPVILEARQXYLQGEEDRLASRADQLEEFVHYVKGKITKVGK"  
complement(17075..17905)  
gene  
Alignment Scores:  
Pred. No.: 7,81e-24 Length: 66714  
Score: 463.00 Matches: 100  
Percent Similarity: 57.45% Conservative: 62  
Best Local Similarity: 35.46% Mismatches: 104  
Query Match: 27.40% Indels: 16  
DB: 1 Gaps: 4  
US-10-724-972A-6352 (1-335) x AF326777 (1-66714)  
QY 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64  
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::  
Db 39269 GTTCAGGACGACACGGCAGCTTTACACTCGAATAAACGCCACACGATGTGTGCTG 39210  
QY 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84  
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::  
Db 39209 GAACTCTGTTCCGCGATGGCTGGCCCGCGGAGCGTCAGCCGATCGTATTGCCGAC 39150  
QY 85 AspAsnLysLysAsnArgIleIleLysProLysArgAspLysIleGlyLysTyrThrSer 104  
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::  
Db 39149 GATAACGATGCAAAACGATCTGCGCGAGCGTGGCCACCTGAAACCGGCGAGTCC 39090  
QY 105 ValGlyThrArgLysGlnProAsnLeuGluGlyIleSerLysLeuLysProAspLeuLe 124  
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::  
Db 39089 GTTCGAAACCGCGCGACGCGAGCTTGAAGGCCATTTCCGCTCTGAAACACACCTGATC 39030  
QY 125 IleAlaAspAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144  
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::  
Db 39029 ATTGCCGACGACGATCGCCATCGGGGGTTTACATCGCTTGCAGCAATTCGCGCGGTA 38970  
QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164  
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::  
Db 38969 CTGCTGTTAAGTCCCGCAACGAAACCTAGCTGAAATTTTGCATCTCGCGGTATCATC 38910  
QY 165 SerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluHisAspLysLys 184  
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::

Db 38909 GCGGAAATGCTGGGTAAAGACGACAGATGTCAGGCACGCTCTGGAACAACATAAGAGAGG 38850  
QY 185 IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204  
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::  
Db 38849 ATGGCGCATGGGCGACCCAG-----CTTCCAAAGGACACGCGTGGCCTTTGGCACA 38796  
QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224  
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::  
Db 38795 TCAGCGGAACAGCAATTCACCTGCATACCTAGGAGACCTGGACCGCGAGCGTGTGGCC 38736  
QY 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu 244  
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::  
Db 38735 TCTCTGGGGCTG-----AACGTTCCCGCTGCGATG 38706  
QY 245 LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263  
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::  
Db 38705 CGGGTTCGTCATGCGGTCCTCGGCACTGCTGCGCGTCAATCTCTGCTGG 38646  
QY 264 MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGluLys 283  
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::  
Db 38645 CTGCTGTTGCCCACTATCGC-----GAAGAGAGCATTTGTTAAACCTGCGCAACA 38595  
QY 284 AspProValTrrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg 303  
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::  
Db 38594 GATCCGCTCTGGCAGATGTTAACCGCGCGCAGAACGACGAGTGTGCTCGTCAGCAGT 38535  
QY 304 AspLeuTrrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuVal 323  
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::  
Db 38534 AACACCTGGCGCGGATGCGCGGTATTTGCTGCAGCGGTATTCGCTGACACGGTA 38475  
QY 324 GluLeu 325  
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::  
Db 38474 AAAATC 38469  
RESULT 35  
AX370197/c 10244 bp DNA linear PAT 16-FEB-2002  
LOCUS Sequence 14 from Patent WO0170776.  
DEFINITION AX370197  
ACCESSION AX370197  
VERSION AX370197.1 GI:18857380  
KEYWORDS Escherichia coli  
SOURCE Escherichia coli  
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Escherichia.  
REFERENCE  
AUTHORS Levy, S.B., Barbosa, T.M. and Alekhan, M.N.  
TITLE Nmr compositions and their methods of use  
JOURNAL Patent: WO 0170776-A 14 27-SEP-2001;  
TRUSTEES OF TUFTS COLLEGE (US)  
FEATURES  
source  
1. 10244  
/organism="Escherichia coli"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:562"  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.39e-24 Length: 10244  
Score: 460.00 Matches: 100  
Percent Similarity: 57.45% Conservative: 62  
Best Local Similarity: 35.46% Mismatches: 104  
Query Match: 27.22% Indels: 16  
DB: 6 Gaps: 4  
US-10-724-972A-6352 (1-335) x AX370197 (1-10244)  
QY 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64  
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::  
Db 7735 GTTCAGGACGACACGCGCAGCTTTACACTCGAATAAACGCCACACGATGTGTGCTG 7676  
QY 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84  
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::

Db 7675 GAACCTCGTTGCGCGATGCGCTGGCCCGCTGGACGTCATCCCGATCGGTATGTCGCGAC 7616  
 QY 85 AspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysThrSer 104  
 Db 7615 GATAACATGCAAAACGATCTCTGCCGAGTGGTGGACCTGAAACCGTGGCAGTCC 7556  
 QY 105 ValGlyThrArgLysGlnProAsnLeuGluLysSerLysLeuLysProAspLeuLeu 124  
 Db 7555 GTCCGACGCGCGCGACGAGCTGGAAGCCATTCGCGCTCTGAAACACAGACCTGATC 7496  
 QY 125 IleAlaAspAsnArgHisLysGlyIleLysLysAspLeuAsnLysIleAlaProThr 144  
 Db 7495 ATTGCCGACAGCAGTCCGCGGCTTACATCCCTTGCAGCAATCGCCCGGTA 7436  
 QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164  
 Db 7435 CTGCTGCTTAAGTCCCGCAACGAAACCTACGCTGAAATTTGCAATCTGCGGCTATCATC 7376  
 QY 165 SerLysAlaLeuGlyLysGluGluGlyLysArgLeuGluGluHisAspLysLys 184  
 Db 7375 GCGCAATGCTGGTAAAGACGAGATCGCGCAGTCTGGAACACATAAGAGAGG 7316  
 QY 185 IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204  
 Db 7315 ATGGCGAGTGGCCAGCCAG-----CTTCCCAAGGGACACGCTGGCTTTGGCACA 7262  
 QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224  
 Db 7261 TCACGCGAAGACGCAATTCAACCTGCATCTCAGGACACCTGGACCGCGTGGCC 7202  
 QY 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu 244  
 Db 7201 TCTCTGGGCTG-----AACCTCCCGCTGGCATG 7172  
 QY 245 LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263  
 Db 7171 GCGGGTGGCTCCATGCGCTCCATCGCTGGAGCAACTCTGCGGTCAATCTGCGTGG 7112  
 QY 264 MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLysLysGluLeuGlyLys 283  
 Db 7111 CTGCTGTTGCCACTATCGC-----GAAGAGACGATGTTAAACGCTGGCAACA 7061  
 QY 284 AspProValTyrLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg 303  
 Db 7060 GATCCGCTCTGCGCATGTTAACCGCGCGAGAGACGAGCTGCTTCGCTGCACAGT 7001  
 QY 304 AspLeuTyrAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuVal 323  
 Db 7000 AACACCTGGCGCGATGCGCGTATTTTGTGCAGAGCGGTATGTCGCGCTGACACGTA 6941  
 QY 324 GluLeu 325  
 Db 6940 AAAATC 6935

RESULT 36  
 U00096\_45/c  
 WPCOMMENT

Sequence split into 47 fragments LOCUS U00096 Accession U00096

Fragment Name	Begin	End
U00096_00	1	110000
U00096_01	100001	210000
U00096_02	200001	310000
U00096_03	300001	410000
U00096_04	400001	510000
U00096_05	500001	610000
U00096_06	600001	710000
U00096_07	700001	810000
U00096_08	800001	910000
U00096_09	900001	1010000
U00096_10	1000001	1110000
U00096_11	1100001	1210000
U00096_12	1200001	1310000
U00096_13	1300001	1410000

U00096\_14 1400001 1510000  
 U00096\_15 1500001 1610000  
 U00096\_16 1600001 1710000  
 U00096\_17 1700001 1810000  
 U00096\_18 1800001 1910000  
 U00096\_19 1900001 2010000  
 U00096\_20 200001 2110000  
 U00096\_21 2100001 2210000  
 U00096\_22 2200001 2310000  
 U00096\_23 2300001 2410000  
 U00096\_24 2400001 2510000  
 U00096\_25 2500001 2610000  
 U00096\_26 2600001 2710000  
 U00096\_27 2700001 2810000  
 U00096\_28 2800001 2910000  
 U00096\_29 2900001 3010000  
 U00096\_30 3000001 3110000  
 U00096\_31 3100001 3210000  
 U00096\_32 3200001 3310000  
 U00096\_33 3300001 3410000  
 U00096\_34 3400001 3510000  
 U00096\_35 3500001 3610000  
 U00096\_36 3600001 3710000  
 U00096\_37 3700001 3810000  
 U00096\_38 3800001 3910000  
 U00096\_39 3900001 4010000  
 U00096\_40 4000001 4110000  
 U00096\_41 4100001 4210000  
 U00096\_42 4200001 4310000  
 U00096\_43 4300001 4410000  
 U00096\_44 4400001 4510000  
 U00096\_45 4500001 4610000  
 U00096\_46 4600001 4639675  
 Continuation (46 of 47) of U00096 from base 4500001 (U00096 Escherichia coli K-12 MG165)

Alignment Scores:  
 Pred. No.: 2,39e-23 Length: 110000  
 Score: 460.00 Matches: 100  
 Percent Similarity: 57.45% Conservative: 62  
 Best Local Similarity: 35.46% Mismatches: 104  
 Query Match: 27.22% Indels: 16  
 DB: 1 Gaps: 4

US-10-724-972A-6352 (1-335) x U00096\_45 (1-110000)

QY 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64  
 Db 12262 GTTCAGGACGAAACACGCGACGCTTTACACTCGAAAAACGCCACACGATTTGGTGTG 12203  
 QY 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84  
 Db 12202 GAATCTGTTGCGCGATGCGCTGCGCCCGCTGACGTCATCCGATCGGTATTGCCGAC 12143  
 QY 85 AspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSer 104  
 Db 12142 GATAACGATGCAAAACGATCTCTCCGAAAGTGGCTGGCAGCTGAAACCGTGGCAGTCC 12083  
 QY 105 ValGlyThrArgLysGlnProAsnLeuGluLysSerLysLysProAspLeuLeu 124  
 Db 12082 GTCCGAAACGCGCGCGACGCGAGCTTGAAGCCATTGCGGCTCTGAAACACGACCTGATC 12023  
 QY 125 IleAlaAspAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144  
 Db 12022 ATTGCCGACAGCAGTCCGATGCGGGGTTTACATCGCTTGCAGCAATCGCCCGGTA 11963  
 QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164  
 Db 11962 CTGCTGCTTAAGTCCCGCAACGAAACCTACGCTGAAATTTGCAATCTGCGGCTATCATC 11903  
 QY 165 SerLysAlaLeuGlyLysGluGluGlyLysArgLeuGluGluHisAspLysLys 184  
 Db 11902 GCGCAATGCTGGTAAAGCGAGATGTCAGGACGCTCTGGAAACCAATAAGAGAGG 11843

```
QY 185 IleGluGluTyrLysGluLeuThrMetAspLysAsnGlnLysValLeuProAlaVal 204
Db 11842 ATGGCGCATGGCCAGCCAG-----CTTCCAAAGGACACGCGCTTGGCCACA 11789
QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224
Db 11788 TCACGGGAACAGCAATTCAACCTGCATCTACGAGACCTGGACCGCAGCGTGGCC 11729
QY 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu 244
Db 11728 TCTCTGGGGCTG-----AACGTTCCCGCTGGCATG 11699
QY 245 LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValLeuProGluAArg 263
Db 11698 GCGGTCGCGTCCATCGCGCTGCATCGCAACCTGCGCGTCAATCTCGCTGG 11639
QY 264 MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLys 283
Db 11638 CTGCTGTTGCCCACTATCC-----GAGAGAGCATTGTTAAACGCTGCAACAA 11588
QY 284 AspProValTrpLysLysLeuAlaValLysAsnGlnArgValAspIleLeuAspArg 303
Db 11587 GATCCGCTCTGGCAGATGTTAACCGCGCGCAGAGCAGCAGTTCCTCGGTGACAGT 11528
QY 304 AspLeuTrpAlaArgSerArgGlyLeuLeuSerSerGluGluMetAlaLysGluLeuVal 323
Db 11527 AACACCTGGCGGGATGCGCGGTATTTTCTGTCAGAGCGTATTCGCGTGCACCGTA 11468
QY 324 GluLeu 325
Db 11467 AAAATC 11462

RESULT 37
ECOW93/c 338534 bp DNA linear BCT 30-JAN-2001
LOCUS Escherichia coli K-12 chromosomal region from 92.8 to 00.1 minutes.
DEFINITION Escherichia coli K-12 chromosomal region from 92.8 to 00.1 minutes.
ACCESSION U14003
VERSION U14003.1 GI:1263172
KEYWORDS
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
Burland, V., Plunkett, G. III, Sofia, H. J., Daniels, D. L. and
Blattner, F. R.
Analysis of the Escherichia coli genome VI: DNA sequence of the
region from 92.8 through 100 minutes
Nucleic Acids Res. 23 (12), 2105-2119 (1995)
95334362
7610040
2 (bases 1 to 338534)
Plunkett, G. I. I.
Direct Submission
Submitted (22-AUG-1994) Guy Plunkett III, Laboratory of Genetics,
University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
Email: ecoliegenetics.wisc.edu Phone: 608-262-2534 Fax:
608-263-7459
On Apr 17, 1996 this sequence version replaced gi:536929.
This sequence was determined as part of the E. coli Genome Project
(Frederick R. Blattner, director) at the University of
Wisconsin-Madison. Supported by award HG00301 from the NIH Human
Genome Project. The entire sequence was independently determined
from E. coli MG1655; overlaps with other sequence determinations
are annotated. The start of this entry overlaps the end of the
entry ECOW89 (U00006) by 1885 bp.
Location/Qualifiers
1..338534
/organism="Escherichia coli"
/mol_type="genomic DNA"
/strain="K-12"
/sub_strain="MG1655"
/db_xref="taxon:562"

FEATURES
source
/imap="92.8 to 100 minutes"
/note="This sequence comprises the following lambda
clones: DD928(EC27-278), DD930(EC30-188),
DD933(EC27-1320), DD935(EC30MM32), DD937(EC27-1070),
DD941(EC17-142), DD945(EC17-8), DD947(EC24A-34),
DD949(EC22-169), DD952(EC27-297), DD953(EC27-409),
DD956(EC17-103), DD958(EC30MM1), DD960(EC30MM2),
DD962(EC21-104), DD965(EC23A-40), DD968(EC30K660A-4pp),
DD970(EC19-202), DD974(EC19-61), DD975(EC18-233),
DD977(EC30MM7), DD980(EC27-1151), DD981(EC18-3),
DD984(EC17-101), DD987(EC18-115), DD990(EC17-136),
DD992(EC18-282), DD995(EC17-115), DD997(EC18-113); M13mp19
or Janus vectors were used for subcloning"
1..1885
/note="1885 bp overlap with end of GenBank Accession
Number U00006 (ECOW89)"
complement(41..398)
/note="ORF_F326 of GenBank Accession Number U00006
(ECOW89)"
/codon_start=1
/transl_table=11
/label=ORF_f132p
/protein_id="AAA96985.1"
/db_xref="GI:1263173"
/translation="MGFTTRVXSESEKPFNPFALFWDKYGTFPILAIIVAFGSLSP
EYFLTNNTIQTIFVQSVTVLIGMGEFFAILVAGIDLSVGAIALSGMVTAKMLAGV
DPLAAMIGVGVGALGALNGCLVNWTL"
misc_structure 277..397
/note="predicted bend of 75.10 degrees"
complement(377..1909)
/note="ORF_f510"
/codon_start=1
/transl_table=11
/label=ORF_f510
/protein_id="AAA96986.1"
/db_xref="GI:536931"
/translation="MATPYISMAGIGKSGFVHALKSVNLTVTPGETHALLGENGAGK
STLMKVLSGHEPTKGTITINNYSNKLHKLAAQIGIYQELSVDELTVLENLY
IGRHLLTKKICGVNIIDWREVRVAAMLRLVGLVDLDEKVALNLSHKQMLEIAKTL
MLDAKVIIMDEPTSSLTKEVDYLFLLMNLQKRGTAIVTISHKLARIICDRTVM
KDSVSGSGIVSDVNDIVLAVGRELQNFAMKVNENSLAHETVFVNVTSRD
KKVRDIFSVCRGELIGFAGLVSGRTELMLNCLPGVDKBRAGELRLKQDI:SPRSPLD
AVKGMAYITTESRRDNGFFNFISIAQNAISRSKDGKYGKAGMLFHEVDQRTAENQ
RELIALKHSVNQITELSGNQOKVLISKWLCCCPVIFDEPTRGIDVGAKEIYK
VMRQLADDGKVLMSSELPEITVCDRIAIVFCERGLTQILTNRDDMSSEIEIMAWALP
QE"
misc_structure 809..931
/note="predicted bend of 74.75 degrees"
1768..1889
/note="predicted bend of 75.87 degrees"
complement(1983..2010)
/note="promoter matrix score of 55; putative; resembles
flh-regulated promoter"
complement(1985..2021)
/standard_name="REP; repetitive extragenic palindromic
element"
/note="contains 1 REP sequence"
complement(2036..2971)
/note="ORF_f311"
/codon_start=1
/transl_table=11
/label=ORF_f311
/protein_id="AAA96987.1"
/db_xref="GI:536932"
/translation="MNKYLKYPFSGTLVGLMLSTSAFAAEYAVVLKTLSPNPFVDMKK
GIEDEAKTLGVSVDIFASPSGDFQSQLPEDLSNKNYKGIAPFAPLSVNLVMPVAR
AWKGIYILNLDKIDMDNLKAGGNVEAFVTDNVAAGKAGASFIIDKLAGEGEVA
IIEGKAGNAGEARRNGATEAFKASQIKLASQPADWDMDRINKALDVATNVLQRNPDIK
AIYCANDTMAMGVAQAVANAGTKGLVVGTDGIPEARRWVEAGQMTATVAQNPAIDG
ATGLKLVDAEKSQKVPILDKAPEFKLVDSILVTQ"
misc_structure 3010..3121
/note="predicted bend of 73.22 degrees"
complement(3030..3953)
```

/note="ORF\_f307"  
/codon\_start=1  
/transl\_table=11  
/label=ORF\_f307  
/protein\_id="AAA96988.1"  
/db\_xref="GI:536933"  
/translation="MASSIQEQENVMSQSFDSALPNGIGLAPVLRMKQSGMTENESR  
IVEWLLKPGNLSKAPAKDVAELAVSEAMVKVSKLGSFGRNLRSALDYPFQSE  
QVLPSEAFDEAPQDVNVKFNFTLTITMEGQSIYVNDIHRARFYQARQDLIGA  
GGSNAICADVQHFELRGVRCQAYPDHIMMSASLQEGDVLVWTHSGRTSDVKAA  
VELAKNGAKILCIHYSHPFIARLADYIICSPAPETPLGRNASARILQTLIDAFF  
VSAUQNEQANIMQKTGAIVDFPFSGALK"  
/complement(3952..3979)  
/note="promoter matrix score of 61; putative"  
4151..4251  
/note="predicted bend of 72.06 degrees"  
4279..4728  
/gene="y7cA"  
4279..4728  
/gene="y7cA"  
/codon\_start=1  
/transl\_table=11  
/label=ORF\_0149  
/protein\_id="AAA96989.1"  
/db\_xref="GI:536934"  
/translation="MKKIAGCDHVGFIKHEIVAHVLRGVEVIDKGTWSSERTDYP  
HYASOVALAVAGEVGGILICTGVGISANKFAGIRAVVCSEPSYASQLSRQHNDF  
NVLAGRSRVGELAKMVDAMUQAQYEGGRQQRVEAITAIEQRN"  
/complement(4520..16191)  
/note="corresponds to GenBank Accession Number D90227  
Number J05260 (SCOPHNAQ) is from E. coli B, not K-12"  
/complement(4739..5146)  
/gene="phnQ"  
/complement(4739..5146)  
/gene="phnQ"  
/note="annotated in GenBank Accession Number J05260  
(SCOPHNAQ), but probably not a gene"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="AAA96990.1"  
/db\_xref="GI:536935"  
/translation="MPAPVPTSSPILIFHSPFILOLTLIKLALAIQVLPAAIFKL  
AAAVILSAPIVAHFTLPVAVRPIFVSFTVRVIVINRLNIANGVVDNIGLGRNVD  
SVSRVKNRSGKRFISQPKETLSLIITHPCK"  
/complement(5272..6030)  
/gene="phnP"  
/complement(5272..6030)  
/gene="phnP"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="AAA96991.1"  
/db\_xref="GI:536936"  
/translation="WSLTLTLTGGAQGVPAWGCCECAACARARSPOYRRQPCGVV  
KFNDAILIDAGLHDLADRSFSGSQFQLLTHYMDHVQGLFPLRWGDDPIPYGPP  
DEQCDDLFRHPLGDLDSHTVEPFFVFDLQGLQVTPILNHSKLTFFGLLETAHSRA  
WLSDTAGPEKTLKFLRNQPVMMDCSHPPRADAPRNHCDLNTVLALANQVIRSPV  
ILTHISHQFADWALMENALPSGFVFGDFGMEIGVA"  
5452..5586  
/note="predicted bend of 75.60 degrees"  
/complement(6032..6466)  
/gene="phnO"  
/complement(6032..6466)  
/gene="phnO"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="AAA96992.1"  
/db\_xref="GI:536937"  
/translation="MPACELRPATQYTDVAYALICELKQAEFDHAFRVGFNANLRD  
PNMRYHALLDGVRVGMGLHVLPHLHVNNWIGIOLBYVMQARGNLVGSKLLAWAE  
EARQAQAEITSLTNVNRHDAHRFVLRGEGYQSHFFTKAL"  
/complement(6453..7010)  
/gene="phnN"

CDS  
complement(6453..7010)  
/gene="phnN"  
/note="HisP-like nucleotide binding protein (ATP-binding  
protein)"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="AAA96993.1"  
/db\_xref="GI:536938"  
/translation="MMKGLIWMCPSSGKDSLLAEILREQTOLLVAHRYITRDASA  
GSNNHIALSQEFTTRAGQNLALSHANGLYIGVGVGIDLWHLHAGDVLVNGRAHL  
PQARARYQSALPVCVQSPEILQRLENRGRENASEINARLARAARYTPQDCHTLANN  
DGSLSQSVDTLLTLIHQKHAHL"  
/complement(7010..8146)  
/gene="phnM"  
/complement(7010..8146)  
Alignment Scores:  
Pred. No.: 9,21e-23 Length: 338534  
Score: 460.00 Matches: 100  
Percent Similarity: 57.45% Conservative: 62  
Best Local Similarity: 35.46% Mismatches: 104  
Query Match: 27.22% Indels: 16  
DB: 1 Gaps: 4  
US-10-724-972A-6352 (1-335) x ECUW93 (1-338534)  
QY 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64  
Db 205073 GTTCAGGACGAACACGGCAGCTTTACACTCGAAAAACGCCACACGGATGTGGTGGCTG 205014  
QY 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84  
Db 205013 GAACCTCTGTCGCGATGCGTGGCCGCGGTGGACGTCATCCCGATCGGTATTGCGGAC 204954  
QY 85 AspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysThrSer 104  
Db 204953 GATAACGATGCAAAACCGCATCTCCCGAAGTGGGTGGCCACCTGAAACCGTGGCAGTCC 204894  
QY 105 ValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLysProLeuIle 124  
Db 204893 GTCCGAAACGCGCGCCGAGCCTGGAAGCCTTGGCGCTCTGAAACACAGACCTGATC 204834  
QY 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144  
Db 204833 ATTGCGACAGCAGTGGCCATCGCGGGGTTTACATCGCTTGCAGCAATCGCGCGGTA 204774  
QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164  
Db 204773 CTGCTGCTTAAGTCCCGCAACGAAACCTACCTGAAATTTGCAATCTGCGGCTATCATC 204714  
QY 165 SerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAspLysLys 184  
Db 204713 GGCGAAATGGTGGGTAAAGCAGAGATGAGGACGCTCTGGAACAACATAAGAGAGG 204654  
QY 185 IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204  
Db 204653 ATGCGCAGTGGGCCACCCAG-----CTTCCCAAGGACACGCGTGGCTTTGGCACA 204600  
QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224  
Db 204599 TCACGCAACAGCAATTCACCTGCATACCTAGAGACCTGACCGGCGGCGTGTGCGCC 204540  
QY 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu 244  
Db 204539 TCTCTGGGCGTG-----AACGTTCCCGCTGCGATG 204510  
QY 245 LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263  
Db 204509 CGCGGTGCGTCCATGCGCGTCCATCGCGCTGAGCACTGTGCGCGGTCAATCTCTGCGCTGG 204450  
QY 264 MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGluLys 283  
Db 204449 CTGCTGTTGCCCACTATCGC-----GAAGAGAGCATTTGTTAAACGCTGGCAACA 204399









Db 15957 GTCTTCTCTGAAACGGTCGCGCGAATGGAAGACAAC-----TTCTCCCTCTAC 16007  
QY 165 SerLysAlaLeuGlyLysGluGluGlyLysLysLysLysLysLysLysLysLysLysLys 184  
Db 16008 GCGAAAGCGTAAACAAAGAGGAAAGGCAAGAGTTTACAGATTTCGATATAAAAA 16067  
QY 185 IleGluGluTyrLysLysGluLleThrMetAspLysAsnGlnLys-----ValLeuPro 202  
Db 16068 GTCAGCGAATTAAAGCGACAAAGCTCGCGACCAAGAAAAGAAACGGGTGTCGTAATCAGA 16127  
QY 203 AlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe 222  
Db 16128 TTTATGCGCGGCAATCAAGATT-----TACTATAAAGATTCTTTCCGGGAATTATC 16181  
QY 223 LeuSerGlnLeuGlyPheLys-----GluAlaLeu-----SerAsp 234  
Db 16182 CTGGATCAGCTCGGATTCAAGTATCGGAAACACAGAAAACCTGTTTGAGAACACAGAC 16241  
QY 235 AspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGlu 254  
Db 16242 GAC-----AAATTTGTTTATGACGGAAAATAAGAA 16274  
QY 255 ThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSerSerAsn 274  
Db 16275 TCGATCCCTGATATGACGAGATGTCCTGTTCTTATTCACATATAAACCTGCCGAA--- 16331  
QY 275 GluProSerLeuLysGluLeuGlu-----LysAspProValTyr 287  
Db 16332 -----AGCAAAAAAGAGCTGAAACATGCGCAAAACGATTGGACAAAGCGATCGCTGTGG 16385  
QY 288 LysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAla 307  
Db 16386 AAAAAATTGAACGCTGTAAAAACCGGAAACGCCCATGAAGTCGACGATCGGATCTGGACG 16445  
QY 308 ArgSerArgGlyLeu 312  
Db 16446 ACGCGCGCGCGGTA 16460  
Search completed: November 9, 2005, 23:37:16  
Job time : 7014 secs

CP000002\_11 1100001 1210000  
CP000002\_12 1200001 1310000  
CP000002\_13 1300001 1410000  
CP000002\_14 1400001 1510000  
CP000002\_15 1500001 1610000  
CP000002\_16 1600001 1710000  
CP000002\_17 1700001 1810000  
CP000002\_18 1800001 1910000  
CP000002\_19 1900001 2010000  
CP000002\_20 2000001 2110000  
CP000002\_21 2100001 2210000  
CP000002\_22 2200001 2310000  
CP000002\_23 2300001 2410000  
CP000002\_24 2400001 2510000  
CP000002\_25 2500001 2610000  
CP000002\_26 2600001 2710000  
CP000002\_27 2700001 2810000  
CP000002\_28 2800001 2910000  
CP000002\_29 2900001 3010000  
CP000002\_30 3000001 3110000  
CP000002\_31 3100001 3210000  
CP000002\_32 3200001 3310000  
CP000002\_33 3300001 3410000  
CP000002\_34 3400001 3510000  
CP000002\_35 3500001 3610000  
CP000002\_36 3600001 3710000  
CP000002\_37 3700001 3810000  
CP000002\_38 3800001 3910000  
CP000002\_39 3900001 4010000  
CP000002\_40 4000001 4110000  
CP000002\_41 4100001 4210000  
CP000002\_42 4200001 4222336  
Continuation (34 of 43) of CP000002 from base 3300001 (CP000002 Bacillus licheniformis D  
Alignment Scores:  
Pred. No.: 5,3e-20 Length: 110000  
Score: 415.50 Matches: 108  
Percent Similarity: 52.62% Conservative: 63  
Best Local Similarity: 33.23% Mismatches: 117  
Query Match: 24.59% Indels: 37  
DB: 1 Gaps: 10  
US-10-724-972A-6352 (1-335) x CP000002\_33 (1-110000)  
QY 8 LeuLysIleLeuSerValIleGlyLeuLeuPheValLeuIleAlaThr-----AlaAla 25  
Db 15537 ATGAGAGATGTCATATTGTTGATTTATCGCTTTGTTAGCGATTTCATTTTCGACGCC 15596  
QY 26 CysGlyAsnAsnSerSerSerSerSerSerLysGluSerLysAspGlyValGluIle 45  
Db 15597 TGCGCGGAGGCAAGGATGACTCTGCAAGGCGCGGAGAAATAAGATACGATCACCGTT 15656  
QY 46 LysHisGluGluGlyThr---ThrLysValProLysHisProLysArgValValValLeu 64  
Db 15657 AAACACGCAATGGGTACCGAAGACCAAGTGCTGCGAGACCCCAAAAGAGTCTGCTGCTG 15716  
QY 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84  
Db 15717 ACAATATGAGNACCGAAGCGCTTTTGGAACTCGGAATCAAAACCTGCGGTGCGATCAAA 15776  
QY 85 AspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSer 104  
Db 15777 TCCTGCAGCGGATCCTTGGTATTAAGCATATTCAAGCGGATATGAAAGGCGTTAAGCAA 15836  
QY 105 ValGlyThrArgLysGlnProAsnLeuGluIleSerLysLeuLysProAspLeuIle 124  
Db 15837 ATCGGACTCGAAGGCGCAACCGAATATCGAAGCGATTGCGAGCTGAAGCCTGATTATATC 15896  
QY 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144  
Db 15897 ATCGGAATAAATGAGACATGAAATAATACGACAAAGCTTGAAAAAATTCCTCCGACT 15956  
QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 9, 2005, 10:23:21 ; Search time 641 Seconds  
(without alignments)  
3093.781 Million cell updates/sec

Title: US-10-724-972A-6352  
Perfect score: 1690  
Sequence: 1 GVSEVRLKILSLVIGLLFVL.....EEMAKELVELSKDKSKDKNK 335

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO spool\_h/US10724972/runat\_03112005\_141859\_29140/app\_query.fasta\_1.519  
-DB=N Geneseq -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPFCU=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10724972 @CGN 1 1 470 @runat\_03112005\_141859\_29140 -NCPUP=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_16Dec04:\*

- 1: geneseqn1980a:\*
- 2: geneseqn1990a:\*
- 3: geneseqn2000a:\*
- 4: geneseqn2001a:\*
- 5: geneseqn2001bs:\*
- 6: geneseqn2002a:\*
- 7: geneseqn2002bs:\*
- 8: geneseqn2003a:\*
- 9: geneseqn2003bs:\*
- 10: geneseqn2003cs:\*
- 11: geneseqn2003ds:\*
- 12: geneseqn2004a:\*
- 13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1690	100.0	1008	13 ADS03285	Ads03285 Staphyloc
2	1671	98.9	993	ACA47084	Aca47084 Prokaryot
3	1242	73.5	1014	AAS54519	Aas54519 Staphyloc
4	1233	73.0	2981	8 ACC48532	Acc48532 Staphyloc
5	1232	72.9	984	8 ACA19774	Aca19774 Prokaryot

6	1214	71.8	1044	8	ACF74986	Acf74986 Staphyloc
7	1181	69.9	927	4	AAS51822	Aas51822 Staphyloc
8	726	43.0	957	4	AAS51612	Aas51612 Staphyloc
9	726	43.0	960	4	AAS54387	Aas54387 Staphyloc
10	724.5	42.9	957	8	ABT15015	Abt15015 Pathogen
11	724.5	42.9	960	8	ACA20079	Aca20079 Prokaryot
12	724.5	42.9	975	8	ACF74415	Acf74415 Staphyloc
13	724.5	42.9	2957	8	ACC48534	Acc48534 Staphyloc
14	721	42.7	2247	2	AAT80398	Aat80398 Staphyloc
15	721	42.7	2247	2	AAT83786	Aat83786 DNA encod
16	679.5	40.2	5718	12	ADF30765	Adf30765 Bacillus
17	644	38.1	525	8	ACA47237	Aca47237 Prokaryot
18	516.5	30.6	897	8	ACA42836	Aca42836 Prokaryot
19	510.5	30.2	801	6	ABK74775	Abk74775 Bacillus
20	476	28.2	909	10	ACF70890	Acf70890 Photorhab
21	476	28.2	25860	11	ADR20887	Adr20887 Photorhab
22	476	28.2	110000	10	ACF67367	Continuation (40 o
23	476	28.2	110000	10	ACF65388	Continuation (9 of
24	460	27.2	900	8	ACA32000	ACA32000 Prokaryot
25	460	27.2	909	4	AAS52735	Aas52735 E. coli D
26	460	27.2	909	4	ACA32751	Aca32751 Prokaryot
27	460	27.2	10244	8	ACA46238	Aa46238 DNA encod
28	454	26.9	891	8	ACA45366	Aca45366 Prokaryot
29	405	24.0	990	8	ABT15032	Abt15032 Pathogen
30	405	24.0	999	2	AAZ22850	Aaz22850 Staphyloc
31	405	24.0	999	2	AAZ19889	Aaz19889 Staphyloc
32	405	24.0	999	10	ADF43555	Adf43555 Staphyloc
33	405	24.0	2940	8	ACC48531	Acc48531 Staphyloc
34	405	24.0	3775	2	AAV74549	Aav74549 Staphyloc
35	391	23.1	972	8	ACA21545	Aca21545 Prokaryot
36	380.5	22.5	944	6	ABK74770	Abk74770 Bacillus
37	375	22.2	2209	5	AAS71379	Aas71379 DNA encod
38	375	22.2	4392	5	AAS73038	Aas73038 DNA encod
39	355	21.0	668	2	AAV74666	Aav74666 Staphyloc
40	350	20.7	2799	5	AAS77335	Aas77335 DNA encod
41	343	20.3	861	8	ACF74610	Acf74610 Staphyloc
42	334	19.8	2115	2	AAV74915	Aav74915 Staphyloc
43	333	19.7	242	2	AAV76867	Aav76867 Staphyloc
44	331	19.6	945	5	AH65340	Aah65340 C glutami
45	331	19.6	349980	5	AH68525	Aah68525 C glutami

# ALIGNMENTS

RESULT 1  
ADS03285  
ID ADS03285 standard; DNA; 1008 BP.  
XX  
AC ADS03285;  
XX  
DT 04-NOV-2004 (first entry)  
XX  
DE Staphylococcus epidermis polynucleotide seqid 2580.  
XX  
KW antibacterial; vaccine; antisense therapy; Staphylococcus epidermidis;  
KW recombinant expression vector; infection; computer readable medium;  
KW computer based system; gene; ds.  
XX  
OS Staphylococcus epidermidis.  
XX  
PN US2004147734-A1.  
XX  
PD 29-JUL-2004.  
XX  
PF 01-DEC-2003; 2003US-00724972.  
XX  
PR 08-NOV-1997; 97US-0064964P.  
PR 13-AUG-1998; 98US-00134001.  
PR 29-NOV-1999; 99US-00450969.  
XX  
(DOUC/) DOUCETTE-STAMM L.  
PA (BUSH/) BUSH D.  
XX

121 ProAspLeuIleAlaAspAsnArgHisLysGlyIleTyrLysAspLeuAsnLys 140  
141 IleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAla 160  
161 PheLysThrIleSerLysAlaLeuGlyLysGluGluGluGlyLysLysArgLeuGluGlu 180  
181 HisAspLysIleGluGluLysLysGluLysGluIleThrMetAspLysAsnGlnLysVal 200  
201 LeuProAlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGly 220  
221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu 240  
241 SerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsn 260  
261 ProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGlu 280  
281 LeuGluLysAspProValTyrLysLysLeuAsnAlaValLysAsnGlnArgValAspIle 300  
301 LeuAspArgAspLeuThrAlaArgSerArgGlyLeuLysSerSerGluGluMetAlaLys 320  
321 TTAGACCGTGACTTATGGCAAGATCACGTGGTTTAAATTTCTCAGAAAGAAATGGCAAA 340  
341 GAACTTGTGAATATCTAAGAAAGATAGTAAAAAGAGTAATATAG 1005

RESULT 2  
ACA47084  
ID ACA47084 standard; DNA; 993 BP.  
AC ACA47084;  
AC ACA47084;  
DT 19-JUN-2003 (first entry)  
DE Prokaryotic essential gene #28741.  
XX Antisense; ds; prokaryotic essential gene; cell proliferation;  
KW drug design; gene.  
XX Staphylococcus epidermidis.  
OS  
XX  
XX WO200277183-A2.  
XX  
XX 03-OCT-2002.  
XX  
XX 21-MAR-2002; 2002WO-US009107.  
XX  
XX 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX  
XX (BLIT-) ELITRA PHARM INC.

PI Doucette-Stamm L, Bush D;  
XX WPI; 2004-580138/56.  
DR P-PSDB; ADS07057.  
XX New isolated polypeptide and encoding nucleic acid derived from  
PT Staphylococcus epidermidis, useful for diagnosing, preventing and/or  
PT treating an S. epidermidis bacterial infection.  
XX  
XX Claim 5; SEQ ID NO 2580; 741pp; English.  
XX The invention describes an isolated nucleic acid comprising a nucleotide  
CC sequence with any of 3772 fully defined nucleotide sequences (SEQ ID NO:  
CC 1-3772) and encoding an Staphylococcus epidermidis polypeptide with any  
CC of 3772 fully defined amino acid sequences (SEQ ID NO: 3772-7544) as  
CC given in the specification. Also described are: a recombinant expression  
CC vector; a cell comprising a recombinant expression vector of (1);  
CC producing an S. epidermidis polypeptide; an isolated nucleic acid  
CC comprising a nucleotide sequence of at least 8 nucleotides in length; a  
CC vaccine composition for prevention or treatment of an S. epidermidis  
CC infection, comprising a nucleic acid cited above and a carrier; treating  
CC a subject for S. epidermidis infection; a recombinant or substantially  
CC pure preparation of an S. epidermidis polypeptide or its fragment; a  
CC vaccine composition for prevention or treatment of an S. epidermidis  
CC infection; detecting the presence of a Staphylococcus nucleic acid in a  
CC sample; a computer readable medium having recorded in it the nucleotide  
CC sequences with SEQ ID NO: 1-3772 or its fragments; a computer based  
CC system for identifying fragments of the Staphylococcus genome of  
CC commercial importance; a computer based system for identifying fragments  
CC of the Staphylococcus plasmids of commercial importance; identifying  
CC commercially important nucleic acid fragments of the Staphylococcus  
CC genome and/or plasmids; and identifying an expression modulating fragment  
CC of the Staphylococcus genome and/or plasmids. The methods and  
CC compositions of the present invention are useful for the diagnosis,  
CC prevention and/or treatment of an Staphylococcus epidermidis bacterial  
CC infection. This sequence encodes a S. epidermis protein of the invention.  
XX  
XX Sequence 1008 BP; 404 A; 129 C; 188 G; 287 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 5, 07e-154 Length: 1008  
Score: 1690.00 Matches: 335  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-10-724-972A-6352 (1-335) x ADS03285 (1-1008)

1 GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeu 20  
1 CGAGTGGAAATCAGTGAGAGGTTTAAAAATTTTAAAGTGTAAATGGCTTATGTTGTTTAA 60  
21 IleAlaThrAlaAlaCysGlyAsnAsnSerSerAsnSerSerLysGluSerLys 40  
61 ATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAGAGTCATCAAAA 120  
41 AspGlyValGluLysHisGluGluGlyThrLysValProLysHisProLysArg 60  
121 GATGGAGTTGAAATCAAGCACGAGAGAGGTACTACGAAAGTACTTAAACACCTAAACGT 180  
61 ValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProVal 80  
181 GTTGTGTTCTTCTGAGTATTCATTTGTTGATGCGCTTGTAGTGTGTTTAAACCTGTT 240  
81 GlyIleAlaAspAsnLysLysAsnArgGlyIleLysProLeuArgAspLysIleGly 100  
241 GGGATACGGGATGATACAAAAAAATCGTATTATTAAACCATTAAGAGTAATAATTGGA 300  
101 LysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluLysSerLysLeuLys 120  
301 AAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAA 360

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX  
DR WPI: 2003-029926/02.  
DR P-PSDB; ABU43214.  
XX  
PT New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
PS Claim 14; SEQ ID NO 34954; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: the sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences)

Sequence 993 BP; 398 A; 128 C; 183 G; 284 T; 0 U; 0 Other;

```

Alignment Scores:
Pred. No.:          3.44e-152          Length:          993
Score:             1671.00             Matches:         331
Percent Similarity: 100.00%             Conservative:    0
Best Local Similarity: 100.00%           Mismatches:     0
Query Match:       98.98%               Indels:         0
DB:                8                   Gaps:           0

```

US-10-724-972A-6352 (1-335) x ACA47084 (1-993)

QY	5	Val	arg	Gly	Leu	lys	lle	Leu	Ser	Val	ile	Gly	Leu	Leu	Phe	Val	Leu	lle	Ala	Thr	Ala	24
DB	1	GT	GAG	AG	CT	T	T	A	A	A	T	T	T	T	T	T	T	T	T	T	T	60
QY	25	Ala	Cys	Gly	Asn	Asn	Ser	Ser	Asn	Ser	Ser	lys	Glu	Ser	Ser	Lys	Asp	Gly	Val	Glu	44	
DB	61	GCA	TG	GGA	AA	TAA	TAG	TT	CAAG	TAA	CTCA	AGTAA	GAGTCA	TCAA	AAAGAT	GGA	TTGAA				120	
QY	45	Ile	Lys	His	Glu	Glu	Gly	Thr	Thr	Lys	Val	Pro	Lys	His	Pro	Lys	Arg	Val	Val	Val	Leu	64
DB	121	AT	CAAG	CAC	CA	GA	AGT	ACT	TAC	CA	AA	AGT	TAA	CAC	CC	CTA	AA	CGT	GTT	GTT	CTT	180
QY	65	Glu	Thr	Ser	Phe	Val	Asp	Ala	Leu	Val	Ala	Leu	Asp	Val	Lys	Pro	Val	Gly	lle	Ala	Asp	84
DB	181	GAG	TAT	TCA	TTG	T	GAT	GCT	TAG	TG	CTT	TAG	ATG	T	TAA	AC	CTG	TGG	GAT	TAG	CGGAT	240
QY	85	Asp	Asn	Lys	Asn	Arg	lle	lle	Lys	Pro	Leu	Arg	Asp	Lys	lle	Gly	Lys	Val	Thr	Ser	104	

Db	241	GATAACAAAAAATCGTATTATTAAACCATTAAGAGATAAAAAATTGGAATAATACACTTCT	300
QY	105	ValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIle	124
Db	301	GTAGGACACAGTAAAGCAACTAACTTAGAAGAAATCAGTAAACTTAAACACAGATTAAAT	360
QY	125	IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr	144
Db	361	ATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAATTGCTCTCTACG	420
QY	145	IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle	164
Db	421	ATTGAACCTGAAAGTTTCGATGGAGATTATAATGAAATATTGATGCTTTTAAAAACAAT	480
QY	165	SerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAspLysLys	184
Db	481	TCAAAAGCTTTAGGTAAAGAAGAAGGTAAAAACGCTTAGAAGAACACGATTAAGAAA	540
QY	185	IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal	204
Db	541	ATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAAATCAAAGGTATTGCCTCGAGTA	600
QY	205	AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer	224
Db	601	GCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCTCTAAGT	660
QY	225	GlnLeuGlyPheLysGluAlaLeuSerAspAspValThrLysGlyLeuSerLysTyrLeu	244
Db	661	CAACTAGTGTTTTAAAGAAGCATTAAAGTGATGTTACTAAAGGTTTAAAGTAAGTATCTT	720
QY	245	LysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMet	264
Db	721	AAAGSACCTTACTTACAATGCAACACTGAACTTTATCTCAAGTGAATCCTGAGCGTATG	780
QY	265	PheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLysAsp	284
Db	781	TTCATAATGACAAACAAAGCAAGTTCTAAACGACCTTCACTTAAAGAACTAGAAAAAGAT	840
QY	285	ProValTyrLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAsp	304
Db	841	CCTGTATGGAGAATAATTAAACGCTGTGAAAAATCAACGTGTGATATATTTAGACCGTGAC	900
QY	305	LeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGlu	324
Db	901	TTATGGCAAGATCACGTGGTTTAATTCTTCAGAGAAATGGCAAAAGAACTTTGTTGAA	960
QY	325	LeuSerLysLysAspSerLysLysAspAsnLys	335
Db	961	TTATCTAAGAAAGATAGTAAAAAAGATAATAAG	993
RESULT 3			
AAS54519			
ID	AAS54519 standard; DNA; 1014 BP.		
XX	AAS54519;		
AC	AC		
DT	13-FEB-2002 (first entry)		
DE	Staphylococcus aureus DNA for cellular proliferation protein #831.		
KW	Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;		
KW	antibacterial; drug design.		
OS	Staphylococcus aureus.		
PN	WO200170955-A2.		
PD	27-SEP-2001.		
PF	21-MAR-2001; 2001WO-US009180.		
PR	21-MAR-2000; 2000US-0191078P.		

319 AATTATCTCTAGGTACACGTAACAGCCAACTTAGAGAAATTTAGTAATTTAAAA 378  
121 ProAspLeuIleAlaAspAenAenArgHisLysGlyIleTyrLysAspLeuAenLys 140  
379 CCGGATTTAATATCGCTAGACATAGACATTAAGGTATTAAATAAGAAATTAACAAA 438  
141 IleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAenIleAspAla 160  
439 ATTGCACCAACATTTATCAATAAGAGTTTTGATGGAGACTACAAACAAACATTAATTCG 498  
161 PheLysThrIleSerLysAlaLeuLysGlyLysGluGluGluLysLysArgLeuGluGlu 180  
499 TTCAAAACAATTCGTAAGCTTTAAATAAGAAAGAGGCGGAAAGCGCTCTTGTCTGAA 558  
181 HisAspLysLysIleGluGluTyrLysLysGluIleThrMetAspLysAenGlnLysVal 200  
559 CATGATAAATTAATCAAAAAGTATAAGATGAAATTAAGTTTGTATAGAAATCAAAAAGTG 618  
201 LeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAenSerTyrValGly 220  
619 CTTCCAGCAGTTGTTGCTAAAGCTGTTTATTAGCACATCCAACTTTCATATGTTGGA 678  
221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu 240  
679 CAATTTTAAACGAACTTGGATTTAAAAATGCATTAAGTATGATGATGTAACAAAAGTTTA 738  
241 SerLysTyrLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsn 260  
739 AGTAAATCTTGAAGACCTTACTTCAATAGTACTGAACATTTAGCTGACTTAAT 798  
261 ProGluArgMetPheIleMetThrAsnLysAlaSerSerAenGluProSerLeuLysGlu 280  
799 CTTCAACGCTGATTAATATGACAGATAATGCTAAAAAGATTTCTGCTGAATTCAGAAG 858  
281 LeuGluLysAspProValTyrLysLysAsnAlaValLysAenGlnArgValAspIle 300  
859 TTACAAGAAGATCCAACTTGGAAAAAGTTGAACCGAGTTAAATAATATCGCGTGATTT 918  
301 LeuAspArgAspLeuTyrAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLys 320  
919 GTTGACCGTGATGTTGGCAAGATCTCGTGCCTTAATTTCTTCTGAAGAAATGCTAAA 978  
321 GluLeuValGluLeuSerLysLysAspSerLys 331  
979 GAACCTTGTGTAATTAATCAAAAAAGACAAAAG 1011  
RESULT 4  
ACC48532/c  
ID ACC48532 standard; DNA; 2981 BP.  
XX AC ACC48532;  
XX AC  
XX 11-AUG-2003 (first entry)  
XX  
XX DE Staphylococcal surface-exposed immunogenic polypeptide DNA.  
XX KW Surface-exposed immunogenic polypeptide; SEIP; siderophore; receptor;  
XX KW antibacterial; vaccine; gene; ds.  
XX OS Staphylococcus aureus.  
XX  
XX FH Location/Qualifiers  
XX FT complement (968. .1951)  
XX FT /\*tag= a  
XX FT /product= "SEIP"  
XX  
XX PN WO2003020875-A2.  
XX  
XX PD 13-MAR-2003.  
XX  
XX PF 17-JUN-2002; 2002WO-US019224.  
XX  
XX PR 17-JUN-2001; 2001US-0298975P.

23-MAY-2000; 2000US-0206848P.  
26-MAY-2000; 2000US-0207272P.  
23-OCT-2000; 2000US-0242578P.  
27-NOV-2000; 2000US-0253625P.  
22-DEC-2000; 2000US-0257931P.  
16-FEB-2001; 2001US-0269308P.  
(BLIT-) ELITRA PHARM INC.  
Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
Yamamoto RT, Xu HH;  
WPI; 2001-611495/70.  
P-PSDB; AAU36660.  
New polynucleotides for the identification and development of  
antibiotics, comprise sequences of antisense nucleic acids.  
Claim 27; SEQ ID NO 8156; 511pp; English.  
The invention relates to antisense inhibitors of genes essential to  
prokaryotic cellular proliferation, their use in identifying the genes,  
their use in the discovery of novel antibiotics, the essential genes  
themselves and the encoded proteins. The prokaryotes used are Escherichia  
coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,  
Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also  
useful for the identification of potential new targets for antibiotic  
development. The antisense nucleic acids can also be used to identify  
proteins used in proliferation, to express these proteins, and to obtain  
antibodies capable of binding to the expressed proteins. The proteins can  
be used to screen compounds in rational drug discovery programmes. The  
antisense nucleic acid sequence is also useful to screen for homologous  
nucleic acids which are required for cell proliferation in a wide variety  
of organisms. The present sequence encodes an essential prokaryotic  
cellular proliferation protein. Note: The sequence data for this patent  
did not form part of the printed specification, but was obtained in  
electronic format directly from WIPO at  
ftp.wipo.int/pub/published\_pct\_sequences  
Sequence 1014 BP; 410 A; 126 C; 187 G; 291 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 1,18e-110 Length: 1014  
Score: 1242.00 Matches: 239  
Percent Similarity: 84.8% Conservatives: 42  
Best Local Similarity: 72.21% Mismatches: 50  
Query Match: 73.4% Indels: 1  
DB: 4 Gaps: 0  
US-10-724-972A-6352 (1-335) x AAS54519 (1-1014)  
QY 1 GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeu 20  
DB 20 GGAGTGGAAACG-ATGAGAGGCTTAAACACTTTTAGTATATGGGATTAATAGTTGCTTA 78  
QY 21 IleAlaThrAlaAlaCysGlyAenAenSerSerSerSerSerSerLysGluSerLys 40  
DB 79 TTTTATGTTGAGCTTGTGTAATACGATAATTCAGTAAATAAGAAATCATCAACTAAA 138  
QY 41 AspGlyValGluLysHisGluGluGlyThrThrLysValProLysHisProLysArg 60  
DB 139 GATACATATTTCGGTAAAGATGAAATGTTGACAGTAAAGTACCTTAAGATGCAAAACGT 198  
QY 61 ValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProVal 80  
DB 199 ATCGTTGTTATGAGTACTCTTTGCGAGATGCATTAGCAGCATTTAGACGTTAAACAGTT 258  
QY 81 GlyIleAlaAspAenLysLysAenArgIleIleLysProLeuArgAspLysIleGly 100  
DB 259 GGTATTGCTGATCATGCTAAGAAAAACGTTATCATTAACAGTTAGAGAAAAATGGG 318  
QY 101 LysTyrThrSerValGlyThrArgLysGlnProAenLeuGluLysLeuLys 120



XX PA (DSQU-) D-SQUARED BIOTECHNOLOGIES INC.  
 XX PI Scott DL;  
 XX DR WPI; 2003-300870/29.  
 XX DR P-PSDB; ABR41865.  
 XX PT Novel surface-exposed immunogenic polypeptide of Staphylococcus aureus  
 PT containing receptors for siderophores or iron-binding ligands, useful for  
 PT producing antibodies effective against Staphylococci infection.  
 XX PS Claim 4; Page 53; 62pp; English.  
 CC The present sequence is that of DNA encoding novel Staphylococcus aureus  
 CC surface-exposed immunogenic polypeptide (SEIP) D2 SA03. To isolate SEIP  
 CC genes, an expression library of S. aureus genomic DNA was screened with  
 CC anti-SEIP antisera. DNA was isolated from clones expressing SEIPs, and  
 CC amino acid sequences were deduced and analysed for conserved and/or  
 CC functional domains. The SEIP encoded by the present DNA sequence has  
 CC sequence homology with the siderophore family of periplasmic binding  
 CC proteins. The invention provides methods for the identification,  
 CC production and recovery of SEIPs. The SEIPs can be used individually, or  
 CC in combination, to produce anti-staphylococcal antibodies useful in  
 CC passive or active immunisation strategies to prevent or contain  
 CC staphylococcal infection. They can also be used to develop diagnostic  
 CC assays  
 XX SQ Sequence 2981 BP; 988 A; 559 C; 386 G; 1048 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 3.36e-109 Length: 2981  
 Score: 1233.00 Matches: 238  
 Percent Similarity: 84.59% Conservative: 42  
 Best Local Similarity: 71.90% Mismatches: 51  
 Query Match: 72.96% Indels: 1  
 DB: 8 Gaps: 0

US-10-724-972A-6352 (1-335) x ACC48532 (1-2981)

QY 1 GluValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeu 20  
 Db 1992 GGAGTGGAAACG-ATGAGAGGCTTAAACACTTTTAGTATATATGGGATTATATGGCTTA 1934  
 QY 21 IleAlaThrAlaAlaCysGlyAsnAsnSerSerSerSerSerSerSerSerSerSerSer 40  
 Db 1933 CTTTATAGTTCGAGTTGCTGTAATACGGATAATTCAGTAAAGAAAGAAATCATCACTAAA 1874  
 QY 41 AspGlyValGluIleLysHisGluGluGlyThrThrLysValProLysHisProLysArg 60  
 Db 1873 GATACATATTCGGTAAACATGAAATGCTACAGTAAAGTACCTAAAGATCAAAACGT 1814  
 QY 61 ValValValLeuGluTyrsSerPheValAspAlaLeuValAlaLeuAspValLysProVal 80  
 Db 1813 ATCGTTGTATTAGAGTACTCATTTTCAGATGCAATAGCAGCATTAGACGTTAAACGATT 1754  
 QY 81 GlyIleAlaAspAsnLysAsnArgIleLysPheProLeuArgAspLysIleGly 100  
 Db 1753 GGTATTGCTGATGCTGTAAGAAAACCTATCATTAACACCATTTAGAGAAAATTCGG 1694  
 QY 101 LysTyThrSerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLysLys 120  
 Db 1693 GATTATATCTCTGTAGGTACAGTAAACAGCCAACTTAGAGGAATATTAGTAATATAA 1634  
 QY 121 ProAspLeuIleIleAlaAspAsnAsnArgHisLysGlyIleTyLysAspLeuAsnLys 140  
 Db 1633 CCGGATTTAATATTCGCTGATAGCATAGATAGATAGATAGATAGATAGATAGATAGAT 1574  
 QY 141 IleAlaProThrIleGluLeuLysSerPheAspGlyAspTyAsnGluAsnIleAspAla 160  
 Db 1573 ATTCACCAACATATATCAATTAAGAGTTTGTATGAGACTACAAACAAAATATTAAATTCG 1514  
 QY 161 PheLysThrIleSerLysAlaLeuGlyLysGluGluGluGlyLysLysLysLysLysLys 180

Db 1513 TTCAAAACAATTCGTAAGCTTTAAATAAGAAAAGAGCGGAAAAACGCTTTCGCTGAG 1454  
 QY 181 HisAspLysLysIleGluGluTyLysLysLysLysLysLysLysLysLysLysLysVal 200  
 Db 1453 CATGATAAATTAATCAATAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 1394  
 QY 201 LeuProAlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyLysValGly 220  
 Db 1393 CTTCCAGCAGTAGTTCGTAAGCTGTTTATTAGCACATCCAACTATTTCATATGTTGA 1334  
 QY 221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAspValThrLysGlyLeu 240  
 Db 1333 CAATTTTAAACCAACTAGGATTTAAATGCAATTAAGTACGATGTAAACAAAGGTTTA 1274  
 QY 241 SerLysTyLysLysGlyProTyLysLeuGlnMetAsnThrGluThrLeuSerGlnValAsn 260  
 Db 1273 AGTAAATATTGAAGAGCACCTTACTTACATTAATAGACACTGAACTATTAGCTGATTAAAT 1214  
 QY 261 ProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGlu 280  
 Db 1213 CCAGACCGTATGATCATTTATGACAGATCATGTAAAAAAGATTCTGCTGAATTCAGAAG 1154  
 QY 281 LeuGluLysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIle 300  
 Db 1153 TTACAGAGAGTCAACATGGAAGAAAGTTGAATGCAGTTAAATAATATCGCTGGATATT 1094  
 QY 301 LeuAspArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLys 320  
 Db 1093 GTTGACCGTATGTTTGGGCAAGATCTCGTGGCTTAATTTCTTCTGAAGAAATGGCTAAA 1034  
 QY 321 GluLeuValGluLeuSerLysLysAspSerLys 331  
 Db 1033 GAACTTGTGTAATTTATCAAAAAAAGAACAAAG 1001

RESULT 5

ACA19774

ID ACA19774 standard; DNA; 984 BP.

AC ACA19774;

DT 19-JUN-2003 (first entry)

DE Prokaryotic essential gene #1431.

KW Antisense; ds; prokaryotic essential gene; cell proliferation;  
 drug design; gene.

OS Staphylococcus aureus.

PN WO200277183-A2.

PD 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

PA (ELIT-) ELITRA PHARM INC.

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlson KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX P-PSDB; ABU15904.

XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.



PT preventing Staphylococcal infection, specifically an infection caused by  
 XX S. aureus, e.g. sepsis.

PS Claim 6; SEQ ID NO 5331; 49pp; English.

CC The invention relates to novel genes and encoded proteins from  
 CC Staphylococcus aureus. A composition comprising the S. aureus protein, a  
 CC nucleic acid encoding the protein, or an antibody to the protein, is  
 CC useful as a pharmaceutical, particularly as a vaccine for treating or  
 CC preventing infection due to Staphylococcus bacteria, specifically an  
 CC infection caused by S. aureus. The composition is particularly useful for  
 CC treating or preventing sepsis in a patient. The composition can also be  
 CC used for diagnostics. The protein is also used in an assay for enzymatic  
 CC studies and as a target for antibiotics. This sequence represents one of  
 CC the novel S. aureus genes of the invention

XX SQ Sequence 1044 BP; 425 A; 128 C; 191 G; 300 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 6.25e-108 Length: 1044  
 Score: 1214.00 Matches: 237  
 Percent Similarity: 84.59% Conservative: 43  
 Best Local Similarity: 71.60% Mismatches: 51  
 Query Match: 71.83% Indels: 2  
 DB: 8 Gaps: 0

US-10-724-972A-6352 (1-335) x ACF74986 (1-1044)

QY 1 GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuPheValLeu 20  
 Db 54 CGAGTGGAAAG-ATGAGAGGTCTAAACCTTTT-AGTATATTGGGATTAATAGTGGCTTTA 111  
 QY 21 IleAlaThrAlaAlaCysGlyAenAenSerSerSerSerSerSerSerSerSerSerLys 40  
 Db 112 CTTTATTGTCAGCTGTGCTAATACGGATAATTCAGGTAAAGAAAGAAATCATCACTAAA 171  
 QY 41 AspGlyValGluIleLysHisGluGluGlyThrLysValProLysHisProLysArg 60  
 Db 172 GATACATATTTCGGTAAAGATGAAATGCTACAGTAAAGTACCTAAAGATCAAAACGT 231  
 QY 61 ValValValLeuGluTyrSerPheValAlaLeuValAlaLeuAspValLysProVal 80  
 Db 232 ATCGTTGTTATTAGTACTCATTTGCGAGATGATGATGAGCATTAAGCGTTAAACCCAGTT 291  
 QY 81 GlyIleAlaAspAspAspLysLysLysLysLysLysLysLysLysLysLysLysLys 100  
 Db 292 GGTATTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 351  
 QY 101 LysTyrThrSerValGlyThrArgLysGlnProAenLeuGluGluLysSerLysLys 120  
 Db 352 GATTATCTCTGTAGGTACACGTAAACAGCCAACTTAGAAGAAATTAGTAAATTTAAA 411  
 QY 121 ProAspLeuIleIleAlaAspAenAenArgHisLysGlyIleTyrLysAspLysLys 140  
 Db 412 CCGGATTATATATCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 471  
 QY 141 IleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAenGluAenIleAspAla 160  
 Db 472 ATTGCACCAATATCATCAATAGAGTTTTCATGAGACTCAACAAATATTAAATTCG 531  
 QY 161 PheLysThrIleSerLysAlaLeuGlyLysGluGluGlyLysLysLysLysLysLys 180  
 Db 532 TTCAAAAACAAATTGCTAAAGCTTTAAATAAGAAAGAAAGCGGAAACGCTTCCTGAG 591  
 QY 181 HisAspLysLysIleGluGluTyrLysLysGluIleThrMetAspLysAenGlnLysVal 200  
 Db 592 CATGATAAATTAATCAATAGATTAAGATGATTAATTAATTTGATAGAAATCAAAAAGTG 651  
 QY 201 LeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAenSerTyrValGly 220  
 Db 652 CTTCCAGCAGTAGTTGCTAAAGCTGGTTTATTAGCATCAACATCAATCATATTGTTGGA 711  
 QY 221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu 240

Db 712 CAATTTTAAACGAACTAGGATTTAAATAATGCATTAAGTACGATGTAACAAAAGGTTTA 771  
 QY 241 SerLysTyrLeuLysGlyProTyrLeuGlnMetAenThrGluThrLeuSerGlnValAen 260  
 Db 772 AGTAAATATTGAAAGGACCTTACTTACAATTAGACACTGAACATTTAGCTGATTTAAAT 831  
 QY 261 ProGluArgMetPheIleMetThrAsnLysAlaSerSerAenGluProSerLeuLysGlu 280  
 Db 832 CCAGAGCGTATGATCATTTATGACAGATCATGCTTAAAGAAAGATTTCTGCTGAATTCAGAAG 891  
 QY 281 LeuGluLysAspProValTyrLysLysLeuAenAlaValLysAenGlnArgValAspIle 300  
 Db 892 TTCAAGAAGATGCAACATGTAAGAAAGTTGAATGCAAGTTAAATAATATCGGTGGATATT 951  
 QY 301 LeuAspArgAspLeuTrpAlaArgSerArgGlyLeuLysSerGluGluMetAlaLys 320  
 Db 952 GTTGACCGTATGTTTGGGCAAGATCTCGTGGCTTAATTTCTTCTGAAGAAATGGCTAAA 1011  
 QY 321 GluLeuValGluLeuSerLysLysAspSerLys 331  
 Db 1012 GAACCTGTTGAATTTATCAAAAAAGAAAGAAAG 1044

RESULT 7

AAS51822

ID AAS51822 standard; DNA; 927 BP.

AC AAS51822;

DT 13-FEB-2002 (first entry)

XX Staphylococcus aureus DNA for cellular proliferation protein #239.

DE Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;

KW antibacterial; drug design.

XX Staphylococcus aureus.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US0009180.

XX 21-MAR-2000; 2000US-0191078P.

XX 23-MAY-2000; 2000US-0206848P.

XX 26-MAY-2000; 2000US-0207272P.

XX 23-OCT-2000; 2000US-0242578P.

XX 27-NOV-2000; 2000US-0253625P.

XX 22-DEC-2000; 2000US-0257931P.

XX 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

XX P-PSDB; AAU33963.

XX New polynucleotides for the identification and development of

XX antibiotics, comprise sequences of antisense nucleic acids.

XX Claim 27; SEQ ID NO 4404; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to

XX prokaryotic cellular proliferation, their use in identifying the genes,

XX themselves and the encoded proteins. The prokaryotes used are Escherichia

XX coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,

XX Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also

XX useful for the identification of potential new targets for antibiotic

XX development. The antisense nucleic acids can also be used to identify

CC proteins used in proliferation, to express these proteins, and to obtain  
CC antibodies capable of binding to the expressed proteins. The proteins can  
CC be used to screen compounds in rational drug discovery programmes. The  
CC antisense nucleic acid sequence is also useful to screen for homologous  
CC nucleic acids which are required for cell proliferation in a wide variety  
CC of organisms. The present sequence encodes an essential prokaryotic  
CC cellular proliferation protein. Note: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 927 BP; 370 A; 120 C; 169 G; 268 T; 0 U; 0 Other;

Alignment Scores: 8.42e-105 Length: 927  
Pred. No.: 1181.00 Matches: 225  
Score: 1181.00  
Percent Similarity: 85.11% Conservative: 38  
Best Local Similarity: 72.82% Mismatches: 46  
Query Match: 69.88% Indels: 0  
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x AAS51822 (1-927)

Qy 16 LeuLeuPheValLeuLeuLeuAlaThrAlaAlaCysGlyAsnAsnSerSerSerSerSer 35  
Db 1 TTAATAGTTGCTTATTTAGTTGCGAGCTTGTGTAATACGGATAATTAAGTAAAAA 60  
Qy 36 LysGluSerSerLysAspGlyValGluLeuLysHisGluGluGlyThrThrLysValPro 55  
Db 61 GAATCATCACTAAGATACCTATTTCCGTAAGAAGATGAAATGCTACAGTAAAAAGTACCT 120  
Qy 56 LysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeu 75  
Db 121 AAGATGCAAAACGATCGTTGTTATTAGTACTCAITTCGAGATGCGATTAAGCAGCATTA 180  
Qy 76 AspValLysProValGlyLeuAlaAspAsnLysLysAsnArgGlyLeuLeuProLeu 95  
Db 181 GACGTTAAACAGTTGGTATGCTGATGATGCTAAGAAACACGATCATTAACACCGATT 240  
Qy 96 ArgAspLysLeuGlyLysThrThrSerValGlyThrArgLysGlnProAsnLeuGluGlu 115  
Db 241 AGAGAAAAATTTGGGAATATATCTCTCTGTAGGTACACGCTAAACACCAACTTAGAGGAA 300  
Qy 116 IleSerLysLeuLysProAspLeuLeuLeuAlaAspAsnAsnArgHisLysGlyLeuTyr 135  
Db 301 ATTAGTAATTAACACCGATTATTAATTCGCTGATGATGATGATGATGATGATGATGAT 360  
Qy 136 LysAspLeuAsnLysLeuAlaProThrLysGluLysSerPheAspGlyAspTyrAsn 155  
Db 361 AAAGAATTAACAAAAATTCACCAACATTTATCAATTAAGAGATTTTGTATGAGACTACAA 420  
Qy 156 GluAsnLysAspAlaPheLysThrLysSerLysAlaLeuGlyLysGluGluGluGlyLys 175  
Db 421 CAAAAATTAATTCGTTCAAAACAATTCGTTAAAGCTTTAAATTAAGAAAAAGACGCGAA 480  
Qy 176 LysArgLeuGluGluHisAspLysLysLysLysGluGluTyrLysLysGluLeuThrMetAsp 195  
Db 481 AAGCGCTTCGCTGAACATGATAATTAATCAAAAGTATAAGATCAAAATTAAGTTGAT 540  
Qy 196 LysAsnGlnLysValLeuProAlaValAlaAlaLysSerGlyLeuLeuAlaHisProSer 215  
Db 541 AGAAATCAAAAAGTGTCTCCAGCAGTTGTTGCTTAAAGCTGTTTATAGCACATCCAAAC 600  
Qy 216 AsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAsp 235  
Db 601 TATTCATATGTTGGCAATTTTAAACGAACCTTGGATTTAAAAATGCAATTAAGTGAT 660  
Qy 236 ValThrLysGlyLeuSerLysThrLysLeuLysGlyProTyrLeuGlnMetAsnThrGluThr 255  
Db 661 GTACAAAAAGGTTTAAGTAATACTTGAAAGGACCTTACTTACAATTAGATACCTGAACAT 720  
Qy 256 LeuSerGlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGlu 275

Db 721 TTAGTGTACTTAATTCCTGAACGCAATGATTATTATGACAGATAATGCTAAAAAGATTCT 780  
Qy 276 ProSerLeuLysGluLeuGluLysAspProValTrpLysLysLeuAsnAlaValLysAsn 295  
Db 781 GCTGAATTCAGAAGATTACAGAAGATCAACTTCGAAAAAGTTGAACGCACTTAAAAAT 840  
Qy 296 GlnArgValAspLysLeuAspArgAspLeuTrpAlaArgSerArgGlyLeuLysSerSer 315  
Db 841 AATCGCGTGGATATTGTTGACCGTGATGTTGGGCAAGATCTCGTGGCTTAATTTCTTCT 900  
Qy 316 GluGluMetAlaLysGluLeuValGlu 324  
Db 901 GAAGAAATGCTAAAGAAGCTTGTGTAA 927

RESULT 8

AAS51612

ID AAS51612 standard; DNA; 957 BP.

XX AAS51612;

XX 13-FEB-2002 (first entry)

XX Staphylococcus aureus DNA for cellular proliferation protein #29.

XX Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;  
antibacterial; drug design.

XX Staphylococcus aureus.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009180.

XX 21-MAR-2000; 2000US-0191078P.

XX 23-MAY-2000; 2000US-0206848P.

XX 26-MAY-2000; 2000US-0207727P.

XX 23-OCT-2000; 2000US-0242578P.

XX 27-NOV-2000; 2000US-0253625P.

XX 22-DEC-2000; 2000US-0257931P.

XX 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
PI Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

XX P-PSDB; AAU33753.

XX New polynucleotides for the identification and development of  
antibiotics, comprise sequences of antisense nucleic acids.

XX Claim 27; SEQ ID NO 4194; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to  
prokaryotic cellular proliferation, their use in identifying the genes,  
themselves and the encoded proteins. The prokaryotes used are Escherichia  
coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,  
Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also  
useful for the identification of potential new targets for antibiotic  
development. The antisense nucleic acids can also be used to identify  
proteins used in proliferation, to express these proteins, and to obtain  
antibodies capable of binding to the expressed proteins. The proteins can  
be used to screen compounds in rational drug discovery programmes. The  
antisense nucleic acid sequence is also useful to screen for homologous  
nucleic acids which are required for cell proliferation in a wide variety  
of organisms. The present sequence encodes an essential prokaryotic  
cellular proliferation protein. Note: The sequence data for this patent  
did not form part of the printed specification, but was obtained in  
electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 957 BP; 382 A; 114 C; 209 G; 252 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 9,596-61 Length: 957  
 Score: 726.00 Matches: 153  
 Percent Similarity: 65.85% Conservativeness: 61  
 Best Local Similarity: 47.08% Mismatches: 99  
 Query Match: 42.96% Indels: 12  
 DB: 4 Gaps: 5

US-10-724-972A-6352 (1-335) x AAS51612 (1-957)

QY 6 AtgGlyLeuLysLysLeuSerValLleGlyLeuLeuPheValLeuLleAlaThraAla 25  
 DB 13 AGGAATATCTTAAATAGTTGTGTATGCTAATCTTCGTT---GTAGCAGTAGCGGT 69  
 QY 26 CysGlyAsnAenSerSerAenSerSerLysGlySerSerLysAspGlyValGluLle 45  
 DB 70 TGTGGTCAAAAGATACT-----GAAGAGAAAACCTGAATGACGACAA 114  
 QY 46 LysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeuGlu 65  
 DB 115 AAAGATGAATTAGCAACTGAAAAAATTAAGAAAAATCCTAAACGTGTTGTTGTTAGAA 174  
 QY 66 TyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyLleAlaAsp 85  
 DB 175 TATAGTTTTCGATTATTAGCAGCATTAGATGAAACCTGTTGTTGTTGATGATGAT 234  
 QY 86 AsnLysLysAsnArgLleLysProLysProLysArgLysLysLysThrSerVal 105  
 DB 235 GCAGCAGTAAATATAACAAGTCAGTAAGATAGAGTGGGGCATATGAATCGGTT 294  
 QY 106 GlyThrArgLysGlnProLysLeuGluLleSerLysLysLysProLysLysLysLle 125  
 DB 295 CGATCTAGACCGCAACCGAATATGAAGTGATGAATTAATTAACCGGATTTGATCAT 354  
 QY 126 AlaAspAsnAenArgHisLysGlyLysLysLysLysLysLysLysLysLysLys 145  
 DB 355 GCAGATGTTAGCAGACATAGAAAAATCAATCAGAAATTAAGCAAAATGCTCCGCAATC 414  
 QY 146 GluLeuLysSerPheAspGlyAspThrAenGluAenLleAspAlaPheLysThrLys 165  
 DB 415 ATGTTAGTACGGTACGGAGATATATATGCAATATTTGATGATGATGATGATGAT 474  
 QY 166 LysAlaLeuGlyLysGluGlyLysLysLysLysLysLysLysLysLysLysLys 185  
 DB 475 AAAGCAGTAGCAAG 534  
 QY 186 GluGluTyrLysLysGluLleThrMetAspLysAenGlnLysValLeuProAlaVal 205  
 DB 535 CGCGAGATTAAGAAAGAAATTAAGCAGATGATGATGATGATGATGATGATGATGAT 594  
 QY 206 AlaLysSerGlyLeuLeuAlaHisProSerAenSerThrValGlyGlnPheLysSer 225  
 DB 595 TCAGAGCAGATGATGTTTATTAATGAAATGAAATGAAATGAAATGAAATGAAAT 654  
 QY 226 LeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLysLysLysLysLys 245  
 DB 655 ATGGGTATTAACCTGAGTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 714  
 QY 246 GlyProTyrLeuGlnMetAsnThrGluThrLysSerGlnValAenProGluArgMetPhe 265  
 DB 715 GGTCCTTATATTTTATTAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 774  
 QY 266 IleMetThrAen---LysAlaSerSerAenGluProSerLysLysGluLeuGlyLys 284  
 DB 775 TTAGCCACTGACGGAAACCGAGCAAAATAGAACGAAATTC-----ATTGAT 822  
 QY 285 Pro---ValTrpLysLysLysAlaValLysAenGlnArgValAspLleLeuAspArg 303  
 DB 823 CCTGCAGTTTGGAAATCATTAAGAGCTGTGAAGATTAACAAAGTTTATGACGTTGACCG 882

QY 304 AspleuTriAlaArgSerArgGlyLeuLysSerSerGluMetAlaLysGluLeuVal 323  
 DB 883 AATAAGTGGTTGAATCAAGGGTATTATCGAAGTGAAGTATGGCAGAGAGATTTGAA 942

QY 324 GluLeuSerLysLys 328

DB 943 AAAATTGCAGAAAA 957

RESULT 9

AAS54387

ID AAS54387 standard; DNA; 960 BP.

XX AAS54387;

XX 13-FEB-2002 (first entry)

DE Staphylococcus aureus DNA for cellular proliferation protein #699.

XX Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;

KW antibacterial; drug design.

XX Staphylococcus aureus.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009180.

XX 21-MAR-2000; 2000US-0191078P.

XX 23-MAY-2000; 2000US-0206848P.

XX 26-MAY-2000; 2000US-0207272P.

XX 23-OCT-2000; 2000US-0242578P.

XX 27-NOV-2000; 2000US-0253623P.

XX 22-DEC-2000; 2000US-0257931P.

XX 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;

PI Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

DR P-PSDB; AAU36528.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids.

XX Claim 27; SEQ ID NO 8024; 51pp; English.

CC The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the genes,  
 CC their use in the discovery of novel antibiotics, the essential genes,  
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia  
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,  
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also  
 CC useful for the identification of potential new targets for antibiotic  
 CC development. The antisense nucleic acids can also be used to identify  
 CC proteins used in proliferation, to express these proteins, and to obtain  
 CC antibodies capable of binding to the expressed proteins. The proteins can  
 CC be used to screen compounds in rational drug discovery programmes. The  
 CC antisense nucleic acid sequence is also useful to screen for homologous  
 CC nucleic acids which are required for cell proliferation in a wide variety  
 CC of organisms. The present sequence encodes an essential prokaryotic  
 CC cellular proliferation protein. Note: The sequence data for this patent  
 CC did not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 960 BP; 385 A; 115 C; 209 G; 251 T; 0 U; 0 Other;

XX Alignment Scores:

Pred. No.: 9,63e-61 Length: 960  
Score: 726.00 Matches: 153  
Conservative: 61  
Best Local Similarity: 65.85%  
Best Local Similarity: 47.08%  
Query Match: 42.96%  
DB: 4 Gaps: 5

US-10-724-972A-6352 (1-335) x AAS54387 (1-960)

QY 6 ArgGlyLeuLysIleLeuSerValIleGlyLeuPheValLeuIleAlaThrAlaA 25  
DB 7 AGGAATATCGTTAAATAGTTGTTTATGCTTAATCTTCGTT---GTAGCAGTAGCGGT 63  
QY 26 CysGlyAsnAsnSerSerAsnSerSerLysSerLysAspGlyValGluIle 45  
DB 64 TGTGGTCAAAAGACTACT-----GAAGAGAAAACCTGAAATGACGACAATA 108  
QY 46 LysHieGluGluGlyThrThrLysValProLysHieProLysArgValValLeuGlu 65  
DB 109 AAAGATGAATTAGGAACCTGAAAATAATTAAGAAAATCCTAAACGTGTTGTTGATTAGAA 168  
QY 66 TyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsp 85  
DB 169 TATAGTTTCTGCTATATTAGCAGCATATAGATATGAACCTGTTGGTATTCAGATGAT 228  
QY 86 AsnLysLysAsnArgIleIleLysProLysArgAspLysIleGlyLysThrSerVal 105  
DB 229 GGCAGCAGTAAATATAACAAAGTCAGTAAGAGATAAGGTTGGGCGCATATGAATCGGTT 288  
QY 106 GlyThrArgLysGlnProAsnLeuGluGluIleSerLysLysLysProAspLeuIle 125  
DB 289 GGATCTAGACCGCAACCAATATGAAGTGTAAATTAATTAACCGGATTTGATCATTA 348  
QY 126 AlaAspAsnAsnArgHieLysGlyIleLysAspLeuAsnLysIleAlaProThrIle 145  
DB 349 GCAGATGTTAGCAGACATATGAATAATCAATCAGAAATTAAGCAAAATTCGTCGACAATC 408  
QY 146 GluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSer 165  
DB 409 ATGTTAGTACGGGTACGGGAGATTATTAATGCAAAATATTGATGCAATTTAAACACAGTCGCT 468  
QY 166 LysAlaLeuGlyLysGluGluGluGlyLysArgLeuGluGluHieAspLysLysIle 185  
DB 469 AAAGCAGTAGCAAGAGAGAAAGAGCGGAGAAACGCTCGGAAAGCATGATAAATATTA 528  
QY 186 GluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAla 205  
DB 529 CGCGAGATTAGAAAGAAATTTGAACAGAGTAGCTTAAACCTGCATTTGCGATTCGATC 588  
QY 206 AlaLysSerGlyLeuLeuAlaHieSerProSerAsnSerTyrValGlyGlnPheLeuSerGln 225  
DB 589 TCAAGACGAGGTATGTTTATTAATAATGAAGATACATTTATGGGACAATTTCTTAATAA 648  
QY 226 LeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLysSerLysTyrLeuLys 245  
DB 649 ATGGGTATTCACCTCGAAGTCAAAAGACAAAACCTACGCATGTTGGTGAAGCGCAAGGT 708  
QY 246 GlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPhe 265  
DB 709 GGTCTCTATTATTATTATTAATAATGAAGAACTTGCCATATCAATCCAAAGTTATGATT 768  
QY 266 IleMetThrAsn---LysAlaSerSerAsnGluProSerLeuLysGluLeuLysAsp 284  
DB 769 TTAGCCACTGACGGAAAACCGCAAAAATAATAGAACGAAATTC-----ATTGAT 816  
QY 285 Pro---ValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg 303  
DB 817 CCTGCAGTTTGGAAATTCATTAAGAGCTGTGAAGATATAACAAAGTTTATGACGTTGACCGA 876  
QY 304 AspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuVal 323  
DB 877 AATAAGTGGTTGAAATCAAGGGGTATATTCGCAAGTGAAAGTATGCGCAGAGATTAGAA 936

QY 324 GluLeuSerLysLys 328  
DB 937 AAAATTGCAGAAAA 951  
RESULT 10  
ABT15015  
ID ABT15015 standard; DNA; 957 BP.  
XX AC ABT15015;  
XX 06-MAR-2003 (first entry)  
XX Pathogen specific antigen related staphylococcal DNA SEQ ID No 301.  
DE DE  
XX Antibacterial; virucide; fungicide; protozoicide; cytostatic; anti-HIV;  
KW hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;  
KW auto-immunity; vaccine; staphylococcal infection; antibody; cancer;  
KW autoimmune disease; HIV; hepatitis; gene; ds.  
XX Staphylococcus sp.  
XX W0200259148-A2.  
XX 01-AUG-2002.  
XX 21-JAN-2002; 2002WO-EP000546.  
XX 26-JAN-2001; 2001AT-00000130.  
XX (CIST-) CISTEM BIOTECHNOLOGIES GMBH.

Meinke A, Nagy E, Von Ahseu U, Klade C, Henics T, Zauner W;  
Minh DB, Vytvytska O, Etz H, Dryla A, Weichhart T, Hafner M;  
Tempelmaier B;  
WPI; 2003-075410/07.  
Identifying, isolating and producing hyperimmune serum-reactive antigens  
from a pathogen, for preparing vaccine or medicament for treating or  
preventing e.g. staphylococcal infections, comprises providing antibody  
preparation.

Example 7; Page 203; 252pp; English.

The invention relates to a novel method for identifying, isolating and  
producing hyperimmune serum-reactive antigens from a pathogen, tumour,  
allergen, a tissue or host prone to auto-immunity, where the antigens are  
used in a vaccine, comprises providing antibody preparation from a plasma  
pool of a type of animal, or individual sera with antibodies against the  
specific pathogen, tumour, allergen, tissue or host prone to auto-  
immunity. The hyperimmune serum-reactive antigens comprising any of the  
62 sequences of 53-2261 amino acids fully defined in the specification,  
or their hyperimmune fragments are useful for the manufacture of a  
pharmaceutical preparation, particularly a vaccine against staphylococcal  
infections or colonisation against S. aureus or S. epidermidis. The  
preparation of antibodies is useful for the manufacture of a medicament  
for treating or preventing staphylococcal infections or colonisation  
against S. aureus or S. epidermidis. The antibody preparations may also  
be used for diagnostic and imaging purposes. Other conditions that can be  
treated include cancer, autoimmune diseases or infections caused by viral  
(e.g. HIV, hepatitis A, B or C), fungal or protozoan pathogens. This  
polynucleotide sequence represents staphylococcal DNA relating to the  
method for identifying and producing pathogen specific antigens of the  
invention

XX Sequence 957 BP; 383 A; 116 C; 210 G; 248 T; 0 U; 0 Other;

Alignment Scores: Pred. No.: 1,34e-60 Length: 957  
Score: 724.50 Matches: 149  
Percent Similarity: 66.67% Conservative: 63  
Best Local Similarity: 46.86% Mismatches: 95  
Query Match: 42.87% Indels: 11

DB: 8 Gaps: 4

US-10-724-972A-6352 (1-335) x ABT15015 (1-957)

Qy 13 ValletGlyLeuLeuPheValLeuLeuAlaAlaCysGlyAsnAsnSerSerSer 32  
 Db 25 GTTGTGTTCTGCTAAATTTAGTTGTAGCTAGCGGTTGCTGCTCAAAAGATACT--- 81  
 Qy 33 AsnSerSerLysGluSerSerLysAspGlyValGluLeuLysHisGluGluGlyThrThr 52  
 Db 82 -----GAAGAGAAACCTGAATGACGACAAATAAAGATGAATAGGAACCTCAA 129  
 Qy 53 LysValProLysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeu 72  
 Db 130 AAAATTAAAGAAATCTTAAACGTTGTTGTTATTAGAAATATAGTTTGTCTGATTATTA 189  
 Qy 73 ValAlaLeuAspValLysProValGlyLeuAlaAlaAspAsnLysLysAsnArgIleile 92  
 Db 190 GCAGCATATAGATGAACCTGTTGTTATTCAGATGTCGACGACCTAAATATAACA 249  
 Qy 93 LysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsn 112  
 Db 250 AAGTCAGTAGAGATAGATGAGGTCATATCAATCGGTTGGATCTAGACCGCAACGAAT 309  
 Qy 113 LeuGluGluLysSerLysLeuLysProAspLeuLeuLeuLeuLeuLeuLeuLeuLeu 132  
 Db 310 ATCGAAGTCATAGTAAATTTAAACCGGATTTGATTCATTCAGATGTTAGCAGACATAAG 369  
 Qy 133 GlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGly 152  
 Db 370 AAAATCAAAATCAGATTGAGCAAAATTTGCTCCGACCAATCATGTTAGTCGGTACGGGA 429  
 Qy 153 AspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 172  
 Db 430 GATTATTAATGCAATATTTGAAGCATTTAAACAGTCGCTAAAGCAGTAGGCAAGAGAAA 489  
 Qy 173 GluGlyLysLeuArgLeuGluGluHisAspLysLysLysLysLysLysLysLysLys 192  
 Db 490 GAAGCGCAGACGCTCTGAAAGACATGATAAATATTAGCGGAGATTAGAAAGAAATTT 549  
 Qy 193 ThrMetAspLysAsnGlnLysValLeuProAlaValAlaAlaLysSerGlyLeuLeuAla 212  
 Db 550 GAACAGAGTACGTTAAATCTGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 609  
 Qy 213 HisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232  
 Db 610 AATAATGAAGATACATATTTATGGGACAAATCTTAATTAATGCTGCTGCTGCTGCTGCT 669  
 Qy 233 SerAspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsn 252  
 Db 670 ACAAAAGACAAACCTACCGATGTTGTTGACCAAGGAGGTTGCTTATATATATTTAAAT 729  
 Qy 253 ThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn---LysAla 271  
 Db 730 AATGAAGAACTTGCCCAATATCATCCAAAGTTATGATTTAGCCACTGACGGAACG 789  
 Qy 272 SerSerAsnGluProSerLeuLysGluLeuGluLysAspPro---ValTrpLysLysLeu 290  
 Db 790 GACAAAAATAGAACCAATTC-----ATTGATCTCGCAGTTTGGAAATCATTA 837  
 Qy 291 AsnAlaValLysAsnGlnArgValAspLeuLeuAspArgAspLeuTrpAlaArgSerArg 310  
 Db 838 AAAGCTGGAAGATTAACAAAGTTTATGACGTTGACCGGAAATAGTGGTTGAAATCAAGG 897  
 Qy 311 GlyLeuLysSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328  
 Db 898 GGGATTATCGCAAGTGAAGTATGCGCAGAGATTTAGAAAAAATTTGCAGAAAAA 951

RESULT 11

ACA20079  
 ID ACA20079 standard; DNA; 960 BP.  
 XX  
 AC ACA20079;

XX  
 DT 19-JUN-2003 (first entry)  
 DE Prokaryotic essential gene #1736.  
 KW Antisense; ds; prokaryotic essential gene; cell proliferation;  
 KW drug design; gene.  
 OS Staphylococcus aureus.  
 PN W0200277183-A2.  
 XX 03-OCT-2002.  
 XX 21-MAR-2002; 2002WO-US009107.  
 XX 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX (ELIT-) ELITRA PHARM INC.  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI: 2003-029926/02.  
 DR P-PSDB; ABU16209.  
 DR  
 DR  
 DR  
 PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 XX  
 PS Claim 14; SEQ ID NO 7949; 1766pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
 CC prokaryotic essential genes. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 960 BP; 385 A; 115 C; 210 G; 250 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 1.34e-60 Length: 960  
 Score: 724.50 Matches: 149

Percent Similarity: 66.67%		Conservative: 63	
Best Local Similarity: 46.86%		Mismatches: 95	
Query Match: 42.87%		Indels: 11	
DB: 8		Gaps: 4	
US-10-724-972A-6352 (1-335) x ACA20079 (1-960)			
QY	13	ValIleGlyLeuLeuPheValLeuLeuAlaThrAlaAlaCysGlyAsnAsnSerSer	32
DB	25	GTGTGTTTATGCTAATCTTAGTTAGTAGCAGTAGCGGGTGTGTCCTCAAAAAGATACT---	81
QY	33	AsnSerSerLysGluSerSerLysAspGlyValGluIleLysHisGluGluGlyThrThr	52
DB	82	-----GAAGAGAAACTGAAATCAGCACAATAAAAGATGAATTAGGAACCTGAA	129
QY	53	LysValProLysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeu	72
DB	130	AAAATTAAAGAAAAATCCTAAACGTTGTTGTATTAGAAATATAGTTTGTCTGATTATTA	189
QY	73	ValAlaLeuAspValLysProValGlyLeuAlaAspAsnLysLysAsnArgIleIle	92
DB	190	GCAGCATTAGATATGAACCTGTGGTATTGCAGATGATGCCAGCCTAAATAATAACA	249
QY	93	LysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsn	112
DB	250	AAGTCAGTAAGATAGATTGGGCGATATGAATCGTTGGATCTAGACCGCAACCGAAT	309
QY	113	LeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsnAsnArgHisLys	132
DB	310	ATGGAAGTGAAGTAAATTAATTAACCGGATTTGATCATTTGCAGATGTTAGCAGACATAAG	369
QY	133	GlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGly	152
DB	370	AAATCAATCAGAAATTCAGCAAAATTCCTCCGCAATCATGTAGTAGCGGTACGGA	429
QY	153	AspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu	172
DB	430	GATTATAATGCAAAATTTTGAAGCATTTAAACAGTCGCTAAAGCAGTAGGCAAGAGAAA	489
QY	173	GluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGluTyrLysLysGluIle	192
DB	490	GNAGCGGAGAGCGTCTGAAAGCATGATATAATATTAGCGGAGATTAGAAAGAAAT	549
QY	193	ThrMetAspLysAsnGlnLysValLeuProAlaValAlaAlaLysSerGlyLeuLeuAla	212
DB	550	GAACAGAGTACGTTAAATCTGCATTTGCAATTCGGTATCTCAAGAGCAGGTATGTTATT	609
QY	213	HisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu	232
DB	610	AATAATGAAGATACATTTATGGGCAATTCCTTAATTAATGCGGTATTCACCTGAAGTC	669
QY	233	SerAspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsn	252
DB	670	ACAAAGACAAAACCTACGCATGTTGGTGAACCAAGGGTGGTCTTATATATATTTAAAT	729
QY	253	ThrGluThrLeuSerGlnValAsnProGluLysArgMetPheIleMetThrAsn---LysAla	271
DB	730	AATGAAGAACTTGCCCAATATCAATCCAAAGTTATGATTTTAGCCACTACCGGAAAAACG	789
QY	272	SerSerAsnGluProSerLeuLysGluLeuGluLysAspPro---ValTrpLysLysLeu	290
DB	790	GACAAAANTAGAACGAAATTC-----ATTGATCCTGCAGTTTGGAAATCATTA	837
QY	291	AsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArg	310
DB	838	AAAGCTGTGAAGATTAACAAGTTTATGACGTTGACCGAAATAAGTGGTTGAAATCAAGG	897
QY	311	GlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys	328
DB	898	GGGATTATCGAAAGTGAAGTATGGCAGAGATTTAGAAAAAATTCGAGAAAAA	951

RESULT 12  
ACF74415

ID	ACF74415 standard; DNA; 975 BP.
XX	ACF74415;
AC	20-NOV-2003 (first entry)
XX	Staphylococcus aureus DNA #2095.
DT	Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
XX	enzymatic assay; antibiotic target; gene; ds.
OS	Staphylococcus aureus.
XX	WO200294868-A2.
PN	28-NOV-2002.
XX	27-MAR-2002; 2002WO-IB002637.
PD	27-MAR-2001; 2001GB-00007661.
XX	(CHIR-) CHIRON SPA.
PA	Masignani V, Mora M, Scarselli M;
XX	WPI; 2003-120786/11.
DR	P-PSDB; ABM72855.
XX	New Staphylococcus aureus protein, useful as a vaccine for treating or
CC	preventing Staphylococcal infection, specifically an infection caused by
CC	S. aureus, e.g. sepsis.
CC	Claim 6; SEQ ID NO 4189; 49pp; English.
CC	The invention relates to novel genes and encoded proteins from
CC	Staphylococcus aureus. A composition comprising the S. aureus protein, a
CC	nucleic acid encoding the protein, or an antibody to the protein, is
CC	useful as a pharmaceutical, particularly as a vaccine for treating or
CC	preventing infection due to Staphylococcus bacteria, specifically an
CC	infection caused by S. aureus. The composition is particularly useful for
CC	treating or preventing sepsis in a patient. The composition can also be
CC	used for diagnostics. The protein is also used in an assay for enzymatic
CC	studies and as a target for antibiotics. This sequence represents one of
CC	the novel S. aureus genes of the invention
XX	Sequence 975 BP; 387 A; 116 C; 216 G; 256 T; 0 U; 0 Other;
SQ	Alignment Scores:
Pred. No.:	1.37e-60 Length: 975
Score:	724.50 Matches: 149
Percent Similarity:	66.67% Conservative: 63
Best Local Similarity:	46.86% Mismatches: 95
Query Match:	42.87% Indels: 11
DB:	8 Gaps: 4
US-10-724-972A-6352 (1-335) x ACF74415 (1-975)	
QY	13 ValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCysGlyAsnAsnSerSerSer 32
DB	43 GTTGTTCATCCTAATTTAGTTAGCAGTAGCGGGTGTGTCCTCAAAAAGATACT--- 99
QY	33 AsnSerSerLysGluSerSerLysAspGlyValGluIleLysHisGluGluGlyThrThr 52
DB	100 -----GAAGAGAAACTGAAATGACGACAATAAAAGATGAATTAGGAACCTGAA 147
QY	53 LysValProLysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeu 72
DB	148 AAAATTAAAGAAAAATCCTAAACGTTGTTGTATTAGAAATATAGTTTGTCTGATTATTA 207
QY	73 ValAlaLeuAspValLysProValGlyIleAlaAspAsnLysLysAsnArgIleIle 92
DB	208 GCAGCATTAGATATGAACCTGTGTTATTCAGATGATGGCAGCAGCAATAATAACA 267



QY 93 LysProLeuArgAspLysLeuIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsn 112  
Db 268 AAGTCAGTACAGATAGATTGGGGCATATGAATCGGTGTGATCTAGACCGCAACCGAAT 327  
QY 113 LeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsnAsnArgHisLys 132  
Db 328 ATGGAAGTGATAGTAATTAATTAACCGGATTTGATCATTCAGATGTTAGCAGACATAAG 387  
QY 133 GlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGly 152  
Db 388 AAAATCAATCAGAAATTCAGCAAAATTCCTCCGCAATCATCTAGTTAGCGGTACGGGA 447  
QY 153 AspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 172  
Db 448 GATTATAATCAATATTGAAGCATTTAAACAGTCGCTAAAGCAGTAGGCAAGAAGAAA 507  
QY 173 GluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGluTyrLysLysGluIle 192  
Db 508 GAAGGCGAGAGCGCTCTGGAAGACGATGATAAATATTAGCGGAGATTAGAAAGAAAT 567  
QY 193 ThrMetAspLysAsnGlnLysValLeuProAlaValAlaAlaLysSerGlyLeuAla 212  
Db 568 GAACAGAGTACGTTAAATCTCGCATTTGCAATTCGATCTCAAGACGAGTATGTTATT 627  
QY 213 HisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232  
Db 628 AATAATGAGATACATTTATGGACATTTCTTAATTAATGGGTATTCAACCTGAAGTC 687  
QY 233 SerAspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsn 252  
Db 688 ACAAGAAGACAAACTACGCATGTTGTGAACGCAAGGCTGCTTATATATATATTAAAT 747  
QY 253 ThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn---LysAla 271  
Db 748 AATGAAGAACTTGCCTAATATCAATCCAAAGTTATGATTTAGCCACTGACGGAAGACG 807  
QY 272 SerSerAsnGluProSerLeuLysGluLeuGluLysAspPro---ValTrpLysLysLeu 290  
Db 808 GACAAAATAGAACGAATTC-----ATTGATCTCTGAGTTTGGAAATCAATTA 855  
QY 291 AsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArg 310  
Db 856 AAAGCTGTGAAGATAACAAAGTTATGACGTTGACCGAATAAGTGTGTAATCAAGG 915  
QY 311 GlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328  
Db 916 GGGATTATCGCAAGTGAAGATGTCAGAGAGATTTAGAAAAAATTCAGAAAAA 969  
RESULT 13  
ACC48534  
ID ACC48534 standard; DNA; 2957 BP.  
XX AC ACC48534;  
XX AC  
XX 11-AUG-2003 (first entry)  
XX DT  
XX DE  
XX DE Staphylococcal surface-exposed immunogenic polypeptide DNA.  
XX KW Surface-exposed immunogenic polypeptide; SEIP; siderophore; receptor;  
XX KW antibacterial; vaccine; gene; ds.  
XX OS Staphylococcus aureus.  
XX FH  
XX FH Key Location/Qualifiers  
XX CDS 983..1960  
XX FT /\*tag= a  
XX FT /partial  
XX FT /product= "SEIP"  
XX FT /note= "No start codon"  
XX PN  
XX PN WO2003020875-A2.  
XX XX  
XX PD 13-MAR-2003.

XX 17-JUN-2002; 2002WO-US019224.  
XX PF  
XX PR  
XX PR 17-JUN-2001; 2001US-0298975P.  
XX PA  
XX PA (DSQU-) D-SQUARED BIOTECHNOLOGIES INC.  
XX PI  
XX PI Scott DL;  
XX WPI; 2003-300870/29.  
XX P-PSDB; ABR41867.

Novel surface-exposed immunogenic polypeptide of *Staphylococcus aureus* containing receptors for siderophores or iron-binding ligands, useful for producing antibodies effective against *Staphylococcal* infection.

Claim 4; Page 55-56; 62pp; English.

The present sequence is that of DNA encoding novel *Staphylococcus aureus* surface-exposed immunogenic polypeptide (SEIP) D2 SA05. To isolate SEIP genes, an expression library of *S. aureus* genomic DNA was screened with anti-SEIP antisera. DNA was isolated from clones expressing SEIPs, and amino acid sequences were deduced and analysed for conserved and/or functional domains. The SEIP encoded by the present DNA sequence has sequence homology with the siderophore family of periplasmic binding proteins. The invention provides methods for the identification, production and recovery of SEIPs. The SEIPs can be used individually, or in combination, to produce anti-staphylococcal antibodies useful in passive or active immunisation strategies to prevent or contain *Staphylococcal* infection. They can also be used to develop diagnostic assays

Sequence 2957 BP; 1160 A; 381 C; 512 G; 904 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.: 5,49e-60 Length: 2957  
Score: 724.50 Matches: 149  
Percent Similarity: 66.67% Conservative: 63  
Best Local Similarity: 46.86% Mismatches: 95  
Query Match: 42.87% Indels: 11  
DB: 8 Gaps: 4

US-10-724-972A-6352 (1-335) x ACC48534 (1-2957)

QY 13 ValIleGlyLeuLeuPheValLeuIleAlaThrAlaLysGlyAsnSerSerSer 32  
Db 1025 GTTGTGTTTATGCTTAATCTTAGTTGATGACGCGGTGTTGGTCAAAAGATACT--- 1081  
QY 33 AsnSerSerLysGluSerSerLysAspGlyValGluIleLysHisGluGlyThrThr 52  
Db 1082 -----GAAGAGAAACTGAATGACGCAATAAAGATGAATTAGGAACCTGAA 1129  
QY 53 LysValProLysHisProLysArgValValLeuGluTyrSerPheValAspAlaLeu 72  
Db 1130 AAAATTAAGAAAAATCCTAAACGTTGTTGTTATTAGATATAGTTTCTGATTATTTA 1189  
QY 73 ValAlaLeuAspValLysProValGlyIleAlaAspAsnLysLysAsnArgIleIle 92  
Db 1190 GCAGCATTAGATATGAACCTGTTGTTGATGATGATGATGATGATGATGATGATGAT 1249  
QY 93 LysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsn 112  
Db 1250 AAGTCAGTAAAGATAGATTGGGGCATATGAATCGGTGTGATCTAGACCGCAACCGAAT 1309  
QY 113 LeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsnAsnArgHisLys 132  
Db 1310 ATGGAAGTGATAGTAATTAATTAACCGGATTTGATCATTCAGATGTTAGCAGACATAAG 1369  
QY 133 GlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGly 152  
Db 1370 AAAATCAATCAGAAATTCAGCAAAATTCCTCCGCAATCATCTAGTTAGCGGTACCGGA 1429  
QY 153 AspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 172

||||| 1430 GATTATAATGCAATATTAAACAGTCTGCTAAAGCAGTACGCAAGAGAA 1489  
173 GluGlyLysArgLeuGluHisAspLysLysLeuGlyLysLysLysLysLysLys 192  
1490 GAAGCGAGAGCGCTCTGAAAGCATGATAAATATTAGCGGAGATTAGAAAGAAAT 1549  
193 ThrMetAspLysAenGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAla 212  
1550 GAACAGAGTACGTTAAATCTGCATTGCAATTCGGTATCTCAAGACGAGTATGTTTAT 1609  
213 HisProSerAsnSerTyValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232  
1610 AATAATGAAGATACATTTATGGGACAAATCTTAATAAATGGTATTCAACCTGAAGTC 1669  
233 SerAspAspValThrLysGlyLeuSerLysTyLysLysGlyProTyLysGlnMetAsn 252  
1670 ACAAAGACAAACTACGATGTTGGTGAACGAGGCTGCTCTATATATATATTTAAAT 1729  
253 ThrGluThrLeuSerGlnValAenProGluArgMetPheLeuMetThrAsn---LysAla 271  
1730 AATGAAGAACTTGCCAAATATCAATCCAAAGATTATGATTTTAGCCACTGACGGAACG 1789  
272 SerSerAenGluProSerLeuLysGluLeuLysLysPro---ValTyrLysLysLeu 290  
1790 GACAAAATAGAACGAAATTC-----ATTGATCTGCAGTTTGGAAATCATTA 1837  
291 AsnAlaValLysAenGlnArgValAspLysLeuAspArgLeuTyrPalaArgSerArg 310  
1838 AAGCTGTGAAGATAACAAGTTTATGACGTTGACCGAATAGTGGTTGAATCAAGG 1897  
311 GlyLeuLeuSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328  
1898 GGGATTATCGCAAGTGAAGTATGCGAAGATTAGAAAAAATTCAGAAAAA 1951

RESULT 14  
AAT80398/C  
ID AAT80398 standard; cDNA; 2247 BP.  
AC AAT80398;  
DT 14-APR-1998 (first entry)  
DE Staphylococcus aureus Gene #4 encoding cDNA sequence 1.  
XX Staphylococcus aureus WCUH 29; antagonist; antibacterial; immunogen;  
KW vaccine; disease; protection; isolation; ss.  
XX Staphylococcus aureus.  
OS WO9731114-A2.  
PN 28-AUG-1997.  
XX 25-FEB-1997; 97WO-GB000524.  
XX 26-FEB-1996; 96GB-00004045.  
XX (SMIK ) SMITHKLINE BEECHAM PLC.  
XX Burnham MKR, Hodgson JE;  
XX WPI, 1997-435166/40.  
XX New Staphylococcus aureus polynucleotide and polypeptide(s) - for  
PT isolating antagonist of the polypeptide(s) useful as anti-bacterials.  
XX Claim 4; Page 30-31; 117pp; English.  
XX The present sequence encodes a novel polypeptide, which is optionally  
CC expressed in NCIMB 40771. The polypeptide, and polynucleotide encoding  
CC it, are derived from Staphylococcus aureus. Cells expressing ligands  
CC binding the polypeptide can be used to isolated candidate compounds that

CC bind and inhibit the activity of the polypeptide. Such compounds can be  
CC used as anti-bacterial compounds. The polypeptide may also be used as an  
CC immunogen to vaccinate an animal for protection against Staphylococcus  
CC aureus caused disease  
XX Sequence 2247 BP; 704 A; 397 C; 307 G; 802 T; 0 U; 37 Other;  
SQ Alignment Scores: 8.49e-60 Length: 2247  
Pred. No.: 721.00 Matches: 169  
Score: 73.48% Conservative: 36  
Percent Similarity: 60.57% Mismatches: 69  
Best Local Similarity: 42.66% Indels: 9  
Query Match: 2 Gaps: 1  
DB: 1  
US-10-724-972a-6352 (1-335) x AAT80398 (1-2247)  
QY 1 GlyValGluSerValArgGlyLeuLysLysLeuSerValLysGlyLeuLeuPheValLeu 20  
Db 831 GGAGTGGGACNGATGAGAGCTTAAAAACCTTTTAGTATATTGGGATTAATATTGGCTTAA 772  
QY 21 IleAlaThrAlaAlaCysGlyAenAenSerSerAen---SerSerLysGluSerSerLy 40  
Db 771 CTTTGTAGTTCGACCTTGTGTATAATCGGATAATCAAGTAAAGAAAGAAATCATCAACTAA 712  
QY 40 sAspGlyVal-GluLysHisGluGluGlyThrThrLysValProLysHisProLysA 60  
Db 711 AGATACTATTTCGGTAAAGATGAAATCGTACAGTAAAGTACCTAAAGATGACCAAC 652  
QY 60 rgValValValLeuGluTyrSerPheValAspAla---LeuValAlaLeuAspValLysP 79  
Db 651 GTATCGTTGTATTAGAGTACTCTTTTGCAGGTTTGCATTTANCAGCATATATCCGCTTANAC 592  
QY 79 ro-ValGlyLysLeuAspAenLysLysAenArgLysLysProLysProLysArgAspLys 98  
Db 591 CAAGTGGTATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 533  
QY 99 IleGlyLysTyThrSerValGlyThrArgLysGlnProAenLeuGluGluLysLys 118  
Db 532 TTGGGGGT-AATACTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 475  
QY 119 LeuLysProAspLeuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 138  
Db 474 TTAANCCGGGATTAATTTATCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 415  
QY 139 AsnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 158  
Db 414 GCCCAGATTCACCAACA-TTATCATTAAGAGTTTGTGATGAGACTACACCCACAATATT 356  
QY 159 AspAlaPheLys-ThrLysSerLysAlaLeuGlyLysGluGluGluGlyLysLysArgLe 178  
Db 355 AATTCGTTCAACAACAATTTGCTAAAGCTTTAAATAAAGANAAAGAGCGNNAAACGCTT 296  
QY 178 uGluGluHisAspLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 198  
Db 295 TGCTGAGCATGATAAATTAATCAATAGTATTAAGATGAATTAATTTGATGAGAAATCA 236  
QY 198 nLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTy 218  
Db 235 CAAAGTGTCTCCAGCAGTAGTTGCTTAAGCTGGTTTATTAGCACATCCACATTTTATA 176  
QY 218 rValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLys 238  
Db 175 TGTGGACAAATTTTAAACGANCATAGGATTTAAAAATGCAATTAAGTGACGATGTAACA 116  
QY 238 sGlyLeuSerLysTyThrLysGlyProTyLysGlnMetAenThrGluThrLeuSerGln 258  
Db 115 AGGTTTAAAGTAAATATTGAAAGGACCTTACTTACAAATTTAGACACTGAAACATTTAGCTGA 56  
QY 258 nValAenProGluAenMetPheLysMetThrAenLysAlaSerSerAsn 274  
Db 55 TTTAAATCCAGAGCGGTATGATCATTTATGACAGATCATGCTGCTAAAAAAGAT 7

RESULT 15	
AAT83786/c	
ID AAT83786 standard; DNA; 2247 BP.	
XX AC AAT83786;	
XX DT 16-JUL-1998 (first entry)	
XX DE DNA encoding 3 Staphylococcus aureus proteins of unknown function.	
XX KW Staphylococcus aureus protein; ribozyme; antisense sequence; control;	
KW vaccine; Staphylococcal infection; bacterial gene expression;	
KW toxic shock syndrome; ss.	
XX OS Staphylococcus aureus.	
XX FH Key Location/Qualifiers	
FT CDS complement(1212..1445)	
FT FT /*tag= a	
CDS complement(1456..1587)	
FT FT /*tag= b	
CDS complement(1705..2016)	
FT FT /*tag= c	
XX WO9730070-A1.	
XX PD 21-AUG-1997.	
XX PF 19-FEB-1997; 97WO-US002318.	
XX PR 20-FEB-1996; 96US-0011888P.	
XX PA (SMIK ) SMITHKLINE BEECHAM CORP.	
XX PI Black MT, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO;	
PI Pratt JM, Reichard RW, Rosenberg M, Ward JM;	
DR WPI: 1997-424969/39.	
DR P-PSDB; AA27819, AA27820, AA27821.	
XX Novel polypeptide(s) from Staphylococcus aureus strain WCUH29 - used to	
PT isolate antimicrobial compounds, and in vaccines against S. aureus	
PT infection.	
XX Claim 9; Page 672-673; 989pp; English.	
CC The present sequence encodes 3 Staphylococcus aureus proteins of unknown	
CC function. The present sequence was isolated from a library of clones of	
CC S. aureus WCUH 29 in Escherichia coli. The DNA sequence can be used in	
CC the construction of ribozymes and antisense sequences to control the	
CC expression of Staphylococcal genes. The DNA sequence is also useful as a	
CC source of regulatory elements for the control of bacterial gene	
CC expression. The encoded protein may be used to produce vaccines to enable	
CC a host to produce specific antibodies with antibacterial action. These	
CC vaccines and antibodies would protect a host against invasion by S.	
CC aureus, and conditions relating to Staphylococcal infection, e.g.	
CC Staphylococcal food poisoning, scaled skin syndrome, and toxic shock	
CC syndrome	
XX SQ Sequence 2247 BP; 704 A; 397 C; 307 G; 802 T; 0 U; 37 Other;	
Alignment Scores:	
Pred. No.:	8.49e-60 Length: 2247
Score:	721.00 Matches: 169
Percent Similarity:	73.48% Conservative: 36
Best Local Similarity:	60.57% Mismatches: 69
Query Match:	42.66% Indels: 9
DB:	2 Gaps: 1

US-10-724-972A-6352 (1-335) x AAT83786 (1-2247)

[illegible]

## RESULT 16

ADDF30765/C	
ID	ADDF30765 standard; DNA; 5718 BP.
XX	
AC	ADF30765;
XX	
DT	12-FEB-2004 (first entry)
XX	
DE	Bacillus subtilis strain MB1510 integration region DNA SEQ ID N
KW	expression library; Gram-positive bacterium host cell;
XX	non-replicating linear integration cassette; gene; Bacillus; ds
OS	Bacillus subtilis.
XX	
FPN	WO2003095658-A1.
XX	
PD	20-NOV-2003.



The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological

CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 897 BP; 293 A; 176 C; 195 G; 233 T; 0 U; 0 Other;

Alignment Scores: Pred. No.: 1.67e-40 Length: 897  
Score: 516.50 Matches: 111  
Percent Similarity: 60.48% Conservative: 65  
Best Local Similarity: 38.14% Mismatches: 98  
Query Match: 30.56% Indels: 17  
DB: 8 Gaps: 4

US-10-724-972a-6352 (1-335) x ACA42836 (1-897)

Qy 43 ValGluLeuLysHisGluGluGlyThrThrLysValProLysHisProLysArgValVal 62  
Db 67 GTTACCGTCAAGATCAAAAGGGGAATTTACCTTAGTAGCGTACCTAAACGGTGTGTT 126  
Qy 63 ValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyLe 82  
Db 127 GCCTTAGAATATCTTATGTGGATGCACCTTGCACAAATTTGGTGTCAGCCCTGGTGTG 186  
Qy 83 AlaAspAspAsnLysLysAsnArgIleLysProLeuArgAspLysIleGlyLysTyr 102  
Db 187 GCAGTACATGATATAAACCCGATTCTGCAGAAAGTACGCGATAGAGTGCAGCCATGG 246  
Qy 103 ThrSerValGlyThrArgLysGlnProAsnLeuGluLysLysLysLysLysProAsp 122  
Db 247 GAATCAGTGGGACCGCTTCTCAACCGAGTTTAGAAGCGGATTTCTGCACTTAACCAT 306  
Qy 123 LeuIleIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLysAsnLysIleAla 142  
Db 307 TTGATCATTGCCGATGATATCCCATCTCTGCGTCTATGAGAACTCAAAAAAATCGCG 366  
Qy 143 ProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspLys 162  
Db 367 CCGACAGTCGTCTTTAATTCGCCCATGAGAACTATCAAGAAACCTTGAAACCGCACAA 426  
Qy 163 ThrIleSerLysAlaLeuGlyLysGluGluGlyLysLysLysArgGluGluHisAsp 182  
Db 427 AAAATCGGTGATTATTAGTAAATCAAAAGAAATGCAAGCGGTATTCACAAACATAAG 486  
Qy 183 LysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuPro 202  
Db 487 CAGGATATTCGGGACATCCCAAA-----ACGTACCGAAGGAAAAAGCGATTATC 540  
Qy 203 AlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe 222  
Db 541 GGTGTTTCAGTGAACCAATTTAATTTATATAATAGGCAATCCTATGCTGTGGCTTAA 600  
Qy 223 LeuSerGlnLeuGlyPheLys-----GluAlaLeuSerAspAspValThrLysGlyLeu 240  
Db 601 GTGGAAGTGCTAGGTATCAAAATGCCAAAGCCCGTCCGATATAC----- 645  
Qy 241 SerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsn 260  
Db 646 -----CAACCTAATGCTTCGTTGGTTTGTAGAACACAGTGGCGGAGAAAG 690

Qy 261 ProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGlu 280  
Db 691 CCTGATCTTGATGCTTAATCCATTATCGT-----GATGAGAGTATTTCAGAGAAA 741  
Qy 281 LeuGluLysAspProValTyrLysLysLeuAsnAlaValLysAsnGlnArgValAspIle 300  
Db 742 TCGGAAATGAAGCCCTTTATGGAAATTTATCCCTCGGTAAATAAATGGTCAAGTCATCTTA 801  
Qy 301 LeuAspArgAspLeuTyrAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLys 320  
Db 802 GCTAATGATTAATTTATGGCAGAGACGCTGCTATTGATCGCGTGAAGTAAATGGCTAA 861  
Qy 321 GluLeuValGluLeuSerLysLysAspSerLys 331  
Db 862 GAAGTCCAGACTTTGTAAACGAAATCCGCCAAA 894  
RESULT 19  
ABK74775  
ID ABK74775 standard; DNA; 801 BP.  
XX ABK74775;  
AC ABK74775;  
DT 13-AUG-2002 (first entry)  
XX 13-AUG-2002 (first entry)  
DE Bacillus licheniformis genomic sequence tag (GST) #2066.  
XX Bacillus licheniformis genomic sequence tag (GST) #2066.  
KW Differential gene expression; genomic sequenced tag; GST;  
KW altered culture condition; environmental stress;  
KW physiological provocation; ds.  
XX Bacillus licheniformis.  
OS Bacillus licheniformis.  
XX W020229113-A2.  
PN W020229113-A2.  
XX 11-APR-2002.  
PD 11-APR-2002.  
XX 05-OCT-2001; 2001WO-US031437.  
PF 05-OCT-2001; 2001WO-US031437.  
XX 06-OCT-2000; 2000US-00680598.  
PR 27-MAR-2001; 2001US-0279526P.  
XX (NOVO ) NOVOZYMES BIOTECH INC.  
XX (NOVO ) NOVOZYMES AS.  
PI Berka R, Clausen IG;  
XX WPI; 2002-416684/44.  
DR WPI; 2002-416684/44.  
XX Monitoring differential expression of several genes in first Bacillus  
PT cell relative to expression of same genes in one or more second Bacillus  
PT cells, by using substrate containing Bacillus genomic sequenced tag  
array.  
XX Claim 4; SEQ ID NO 2066; 200pp; English.  
PS The invention describes a method of monitoring differential expression of  
XX genes in a first Bacillus cell relative to expression of the genes in  
XX other Bacillus cells, comprising hybridising labelled nucleic acid probes  
XX isolated from Bacillus cells to a substrate containing array of Bacillus  
XX genomic sequenced tags (GST), examining the array, and determining  
XX relative gene expression by an observed hybridisation reporter signal of  
XX a spot in the array. The method is useful for measuring the expression of  
XX genes in a first Bacillus cell relative to expression of the same genes  
XX in one or more second Bacillus cells. The method is useful for monitoring  
XX global expression of several genes from a Bacillus cell, discovering new  
XX genes, identifying possible functions of unknown open reading frames and  
XX monitoring gene copy number variation and stability. Monitoring of the way  
XX in which Bacillus cells adapt to changes in culture conditions,  
XX environmental stress or other physiological provocation. Extensive follow  
XX up characterisation is unnecessary, when one spot on an array equals one  
XX gene or one open reading frame, since sequence information is available.

CC This sequence represents a genomic sequence tag (GST) used in the method  
CC of the invention. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 801 BP: 257 A; 166 C; 204 G; 174 T; 0 U; 0 Other:

**Alignment Scores:**

Pred No.:	5.54e-40	Length:	801
Score:	510.50	Matches:	108
Percent Similarity:	62.88%	Conservative:	58
Best Local Similarity:	40.91%	Mismatches:	95
Query Match:	30.21%	Indels:	3
DB:	6	Gaps:	3

US-10-724-972A-6352 (1-335) x ABK74775 (1-801)

Qy	60	ArgValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysPro	79
Db	3	AAAGTTGTGGTCTTCGAAATGGGATTTATTGATCGCTCTTGATCGGGTATTAAAGCCT	62
Qy	80	ValGlyLeAlaAspAspLeuLysLysAsnArgLleleLysProLeuArgAspLysIle	99
Db	63	GTCCGGAAATTGCGACACGCGCAACCTTAACTTTATTAAACGAGAGGTTCAGGGGAAAATC	122
Qy	100	GlyLysThrSerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeu	119
Db	123	AAAGGGTATATTCACTCGGTTCGCGGCCACCAAGCTTGAAAAAATTTGCTCTCTTA	182
Qy	120	LysProAspLeulleleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsn	139
Db	183	AAGCCCGATTTAATTATTCGGATTTCGACGAGCACAGCGGCGTCTATGATAAGCTGTGC	242
Qy	140	LysIleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAsp	159
Db	243	AAATTTGCGCCACAAATCCGGCTCAAAAACTTGAATTCGCGATTATCAGGACACGCTTGAT	302
Qy	160	AlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGluGlyLysLysArgLeuGlu	179
Db	303	GCATCTCTTACGATTTCGGAAGCCGTCGCGCAAGAAAGCGCAATGCAGAAAAATTTGGCT	362
Qy	180	GluHisAspLysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLys	199
Db	363	GAACATAACAAAGAACTGGATGAATTTGAACACAGAAATTCGCG---AGCCGGAAACACAGAGC	419
Qy	200	ValLeuProAlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrVal	219
Db	420	ATTTCTTCTGCTCGGGAACACAAATCAGGAAATTCACGTCGCGATGAAACTTTTTCACG	479
Qy	220	GlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAspValThrLysGly	239
Db	480	TCTCAGCTTCTGACGAAATCGGCTATACATACGGTGTCCGAGACAGC---GGCAAGGC	536
Qy	240	LeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnVal	259
Db	537	GATCCCGAAACGGTGAATCCGTCATATTTAAATGACGCTCGAACAGCTGCTTCAGAAA	596
Qy	260	AsnProGluArgMetPheIleMetThr---AsnLysAlaSerSerAsnGluProSerLeu	278
Db	597	GATCCCGACGTTATTGTCTCTGATGCGGAGAAAAAGATAAAGTCGATGAACCGGAAA	656
Qy	279	LysGluLeuGluLysAspProValTrrpLysLysLeuAsnAlaValLysAsnGlnArgVal	298
Db	657	AGACCGATCGAAAAAGATCTCTTTGGAAAAAGCTCAGCGCAGTCAAAACCGCAAGGT	716
Qy	299	AspileLeuAspArgAspLeuTrrpAlaArgSerArgGlyLeuIleSerSerGluGluMet	318
Db	717	TATCAACCCGACAGATTCCGCTGGTCCGACGACGATTTGACGGGAGCAGATGAGCTG	776
Qy	319	AlaLysGluLeu	322
Db	777	ATGGATGAAATC	788

RESULT 20

ACF70890	
ID	ACF70890 standard; DNA; 909 BP.
XX	
AC	ACF70890;
AC	
DT	20-NOV-2003 (first entry)
XX	
DE	Photorhabdus luminescens nucleotide sequence #9357.
XX	
KW	Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
KW	detection; food; gene expression; plant; animal; microorganism; toxin;
KW	antibiotic; biopesticide; virulence factor; disease model; plague;
KW	whooping cough; gene; ds.
XX	
OS	Photorhabdus luminescens.
XX	
FN	WO200294867-A2.
XX	
PD	28-NOV-2002.
XX	
PF	07-FEB-2002; 2002WO-IB003040.
XX	
PR	07-FEB-2001; 2001FR-00001659.
XX	
XX	(INSP ) INST PASTEUR.
PA	(CNRS ) CNRS CENT NAT RECH SCI.
PA	
PI	Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
PI	Buchrieser C;
XX	
DR	WPI; 2003-148459/14.
XX	
XX	Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
PT	useful e.g. as therapeutic antimicrobials and agricultural pesticides.
PT	
XX	Claim 2; SEQ ID NO 9357; 1205pp; French.
XX	
XX	The invention relates to the isolation of genes and their encoded
CC	proteins from Photorhabdus luminescens. The isolated sequences are
CC	sources of probes and primers for detecting the genome of P. luminescens
CC	and related species; to study polymorphisms; for gene analysis and for
CC	detection/amplification of the genes. Antibodies (Ab) raised against the
CC	polypeptides encoded by the genes are used for detection/identification
CC	of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
CC	carry a gene-containing vector are used to select compounds that
CC	modulate, regulate, induce or inhibit expression of the genes in plants,
CC	animals or microorganisms other than P. luminescens and are able to alter
CC	response or sensitivity to toxins and antibiotics produced by P.
CC	luminescens. Cells transformed to express the genes are useful for
CC	recombinant production of the proteins, particularly toxins and
CC	antibacterials useful as insecticides, bactericides and fungicides. The
CC	genes, proteins, vectors containing the genes and Ab are also useful
CC	therapeutically (to treat microbial infection by bacteria or fungi that
CC	are sensitive to P. luminescens-encoded toxins or antibiotics) and as
CC	biopesticides. Other uses of the genes and the proteins are as virulence
CC	factors and for identifying targets of human diseases for which P.
CC	luminescens is a model (particularly plague and whooping cough). This
CC	sequence represents one of the isolated P. luminescens genes
XX	
SQ	Sequence 909 BP; 230 A; 258 C; 225 G; 196 T; 0 U; 0 Other:

Alignment Scores:

Pred. No.:	1.42e-36	Length:	909
Score:	476.00	Matches:	102
Percent Similarity:	57.24%	Conservative:	60
Best Local Similarity:	36.04%	Mismatches:	107
Query Match:	28.17%	Indels:	14
DB:	10	Gaps:	3

US-10-724-972A-6352 (1-335) x ACF70890 (1-909)

Qy 43 ValGluileLysHisGluGluGlyThrThrLysValProLysHisProLysArgValVal 63

Db 70 GTACGGTACAGATGAGCGGCGAGTTTACCTCAATACCGTTCCCGAGCGGTAGTC 129  
Qy 63 ValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIle 82  
Db 130 GTGCTGGATTGTCAATTCGCCGATGATGGCGCCATCAATATCAGTCGGTAGGTATT 189  
Qy 83 AlaAspAsnLysLysAsnArgIleLysProLeuArgAspLysIleGlyLysTyr 102  
Db 190 GCCGATGATAACGATCCGCGAGCGCATCTCCAGCATGTTTCGCCAGCGTATTAAAGCCCTCG 249  
Qy 103 ThrSerValGlyThrArgLysGlnProAsnLeuGluGluLysSerLysLysLeuLysProAsp 122  
Db 250 CAATCAACCGGCGACTCGCGCCCAACCAAGCCTTGAAGCCATCAGCGCGTGAACCTGAT 309  
Qy 123 LeuIleAlaAspAsnArgHisLysGlyIleTyrLysAspLysLeuLysIleAla 142  
Db 310 CTCATCATTCGACAGCAGTCAACCCAGCGCGGATTTATCGGGCACTCAAGGGCATCGCT 369  
Qy 143 ProThrIleGluLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLys 162  
Db 370 CCGTCTGCTGCTAAATCCCGAACCAACCTACGAGAGAACCTGCAATCTGCCGCG 429  
Qy 163 ThrIleSerLysAlaLeuGlyLysGluGluGluLysArgLeuGluGluHisAsp 182  
Db 430 ATCATCGGCAAGTCATGCGGAAGAGAGTGAGATGCGAAACCGCTGGCAGACCGCT 489  
Qy 183 LysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuPro 202  
Db 490 GAACGCGATTAAGGCTATGCCAGCAG-----TTGCCACAAGCGCTCAGTGTGTTT 543  
Qy 203 AlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe 222  
Db 544 GGCACCTCGCGCAACAGCAATTTAACTGCACCTCCAGCGATACCTATACCGCAGCGTC 603  
Qy 223 LeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLys 242  
Db 604 TTAACAGCATTTAGGGCTAAAGTTCCCGCATCATGTTAAAT----- 642  
Qy 243 TyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGlu 262  
Db 643 -----CATGGCGCAATGTTCTCTCAACCTGGACATTTACTGGCACTCAATCTGTAC 696  
Qy 263 ArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGlu 282  
Db 697 TGGCTTATCGTGGCCCACTACTCGT-----CAGGAAAGCATCGTGAAACGCTGGCAG 747  
Qy 283 LysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnAtgValAspIleLeuAsp 302  
Db 748 CAGGATACTCTGTGCAGATGATGACGACGACGACGACGACGACGACGACGACGACGAC 807  
Qy 303 ArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluLeuMetAlaLysGluLeu 322  
Db 808 AGCAACACCTGGGCTCGTATGCGCGGATATTTGGCGCTGAACGCACTTGGCAGCATGCG 867  
Qy 323 ValGluLeu 325  
Db 868 GTAAAAATC 876  
RESULT 21  
ID ADR20887/c  
ADR20887 standard; DNA; 25860 BP.  
AC ADR20887;  
XX  
XX  
XX 07-OCT-2004 (first entry)  
DE Photorhabdus luminescens genomic DNA encoding two insecticide proteins.  
XX insecticide protein; transgenic plant; insect-resistant plant; wheat;  
KW maize; gene; ds.  
XX Photorhabdus luminescens subsp. laumondii.

XX PH Key Location/Qualifiers  
CDS 20872..21306  
FT /\*tag= a  
FT /product= "Photorhabdus luminescens insecticide protein  
FT #1"  
FT 21345..22598  
FT /\*tag= b  
FT /product= "Photorhabdus luminescens insecticide protein  
FT #2"  
XX WO2003087377-A1.  
PN  
XX 23-OCT-2003.  
XX 17-APR-2003; 2003WO-FR001239.  
XX 17-APR-2002; 2002FR-00004798.  
XX (INSP ) INST PASTEUR.  
PA (CNRS ) CNRS CENT NAT RECH SCI.  
PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
XX WPI; 2003-833739/77.  
DR P-PSDB; ADR20889, ADR20891.  
XX  
XX New nucleic acid from Photorhabdus luminescens, useful for producing  
PT insecticidal polypeptides active against Lepidoptera, and for producing  
PT insect resistant transgenic plants.  
XX Claim 22; SEQ ID NO 1; 44pp; French.  
XX  
XX The invention comprises the amino acid and coding sequences of proteins  
CC from Photorhabdus luminescens which are active against insects. The  
CC proteins of the invention are toxic to Lepidoptera, specifically the  
CC genera Plutella, Heliothis, Helicoverpa, Spodoptera and Ostrinia. The DNA  
CC and protein sequences of the invention are useful in the preparation of  
CC transgenic, insect-resistant plants, specifically wheat and maize. The  
CC present Photorhabdus luminescens genomic DNA contains the coding sequence  
CC for the two insecticide proteins of the invention.  
XX  
XX Sequence 25860 BP; 7315 A; 5392 C; 6047 G; 7106 T; 0 U; 0 Other;  
SQ  
Alignment Scores:  
Pred. No.: 9.31e-35 Length: 25860  
Score: 476.00 Matches: 102  
Percent Similarity: 57.24% Conservative: 60  
Best Local Similarity: 36.04% Mismatches: 107  
Query Match: 28.17% Indels: 14  
DB: 11 Gaps: 3  
US-10-724-972A-6352 (1-335) x ADR20887 (1-25860)  
Qy 43 ValGluIleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValVal 62  
Db 8507 GTCCGGTACAGATGAGCGGCGATTTTACCCTCAATACCGTTCCCGAGCGGTAGTC 8448  
Qy 63 ValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIle 82  
Db 8447 GTGCTGGAAATGTCATTTGCCGATGATGGCGCCATCAATATCAGTCGGTAGGTATT 8388  
Qy 83 AlaAspAsnLysLysAsnArgIleLysProLeuArgAspLysIleGlyLysTyr 102  
Db 8387 GCCGATGATAACGATCCGCGAGCGCATTTCTACCGATTTTCGCCAGCGTATTAAAGCCCTCG 8328  
Qy 103 ThrSerValGlyThrArgLysGlnProAsnLeuGluGluLysSerLysLysLeuLysProAsp 122  
Db 8327 CAATCAACCGGCGACTCGCGCCCAACCAAGCCTTGAAGCCATCAGCGCGTGAACCTGAT 8268  
Qy 123 LeuIleAlaAspAsnArgHisLysGlyIleTyrLysAspLysLeuLysIleAla 142  
Db 8267 CTCATCATTCGACAGCAGTCAACGCCACGCGGGGATTTATCGGCACTCAAGGGCATCGCT 8208



143 ProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLys 162  
163 ThrIleSerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAsp 182  
8147 ATCATCGCAAGTCATGGGAAAGACAGTATGATGACGAAACCGCTGGCAGACACCGT 8088  
183 LysLysIleGluGluTyrLysLysLysLysLysLysLysLysLysLysLysLysLys 202  
8087 GAACGCATGAAGGCTATGCCAGCCAG-----TTGCCACAAGCGCTCAGTGTGTTT 8034  
203 AlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe 222  
8033 GGCACCTCGCGCAACACGCAATTAACCTGCACCTCCAGCGATACCTATACCGGAGCGTC 7974  
223 LeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLys 242  
7973 TTAACAGCATTAGGCTAAAGTTCCTCCGATCAGTTAA-----7935  
243 TyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGlu 262  
7934 -----CATCGCGCAATGGTTCTCTCAACCTGGACAGTACTGGCACTCAATCCTGAC 7881  
263 ArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGlu 282  
7880 TGCTTATCGGCCCACTATCGT-----CAGGAAAGCATCGTGAACGCTGGCAG 7830  
283 LysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAsp 302  
7829 CAGGATACCTGTGGCAGATGATGACAGCAGCACACAGCATCAGATGACGGGTAGAC 7770  
303 ArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeu 322  
7769 AGCAACACCTGGGCTCGTATCGCGGATATTTCGCGTGAAGCATTCGGCAGCATGCG 7710  
323 ValGluLeu 325  
7709 GTAAAAATC 7701

RESULT 22  
ACF67367\_39/c

Continuation (40 of 57) of ACF67367 from base 3900001 (Photobacterium luminescens nucleotid  
WP Sequence split into 57 fragments LOCUS ACF67367 Accession ACF67367

Fragment Name	Begin	End
WP ACF67367_00	1	110000
WP ACF67367_01	100001	210000
WP ACF67367_02	200001	310000
WP ACF67367_03	300001	410000
WP ACF67367_04	400001	510000
WP ACF67367_05	500001	610000
WP ACF67367_06	600001	710000
WP ACF67367_07	700001	810000
WP ACF67367_08	800001	910000
WP ACF67367_09	900001	1010000
WP ACF67367_10	1000001	1110000
WP ACF67367_11	1100001	1210000
WP ACF67367_12	1200001	1310000
WP ACF67367_13	1300001	1410000
WP ACF67367_14	1400001	1510000
WP ACF67367_15	1500001	1610000
WP ACF67367_16	1600001	1710000
WP ACF67367_17	1700001	1810000
WP ACF67367_18	1800001	1910000
WP ACF67367_19	1900001	2010000
WP ACF67367_20	2000001	2110000
WP ACF67367_21	2100001	2210000
WP ACF67367_22	2200001	2310000
WP ACF67367_23	2300001	2410000
WP ACF67367_24	2400001	2510000
WP ACF67367_25	2500001	2610000
WP ACF67367_26	2600001	2710000
WP ACF67367_27	2700001	2810000

WP ACF67367\_28 2800001 2910000  
WP ACF67367\_29 2900001 3010000  
WP ACF67367\_30 300001 3110000  
WP ACF67367\_31 3100001 3210000  
WP ACF67367\_32 3200001 3310000  
WP ACF67367\_33 3300001 3410000  
WP ACF67367\_34 3400001 3510000  
WP ACF67367\_35 3500001 3610000  
WP ACF67367\_36 3600001 3710000  
WP ACF67367\_37 3700001 3810000  
WP ACF67367\_38 3800001 3910000  
WP ACF67367\_39 3900001 4010000  
WP ACF67367\_40 4000001 4110000  
WP ACF67367\_41 4100001 4200001  
WP ACF67367\_42 4200001 4310000  
WP ACF67367\_43 4300001 4410000  
WP ACF67367\_44 4400001 4510000  
WP ACF67367\_45 4500001 4610000  
WP ACF67367\_46 4600001 4710000  
WP ACF67367\_47 4700001 4810000  
WP ACF67367\_48 4800001 4910000  
WP ACF67367\_49 4900001 5010000  
WP ACF67367\_50 5000001 5110000  
WP ACF67367\_51 5100001 5210000  
WP ACF67367\_52 5200001 5310000  
WP ACF67367\_53 5300001 5410000  
WP ACF67367\_54 5400001 5510000  
WP ACF67367\_55 5500001 5610000  
WP ACF67367\_56 5600001 5648894

Alignment Scores:

Pred. No.: 5,69e-34 Length: 110000  
Score: 476.00 Matches: 102  
Percent Similarity: 57.24% Conservative: 60  
Best Local Similarity: 36.04% Mismatches: 107  
Query Match: 28.17% Indels: 14  
DB: 10 Gaps: 3

US-10-724-972A-6352 (1-335) x ACF67367\_39 (1-110000)

QY 43 ValGluIleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValVal 62  
Db 95117 GTCACGGTACAGATGAGAGCGGAGTTTACCTCAATACCGTTCCCGAGCGGTAGTC 95058  
QY 63 ValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyLe 82  
Db 95057 GTGCTGGAAATTGTCATTTCCGATGTCATGGCGCCATCAATATCAGTCCGCTAGTATT 94998  
QY 83 AlaAspAsnLysLysAsnArgIleLysProLeuArgAspLysIleGlyLysTyr 102  
Db 94997 GCCGATGATACGATCCCGACGCGATCTCCCGATGTTCCGCGATTTAAGCCCTGG 94938  
QY 103 ThrSerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLysProAsp 122  
Db 94937 CAATCAACCGGCACCTCGCGCCCAACCAAGCCTTGAAGCCATCAGCGCGCTGAAACCTGAT 94878  
QY 123 LeuIleIleAlaAspAsnArgHisLysGlyLysLysLysLysLysLysLysLysAla 142  
Db 94877 CTCATCATTTGAGACAGTCAACGCGCGGGATTTATCGGCGACTCAAGGCGATCGCT 94818  
QY 143 ProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLys 162  
Db 94817 CCTGCTCTGCTGCTAAATCCCGCAAGAAACCTACGAAAGAGAACCTGCAATCTGCCCGC 94758  
QY 163 ThrIleSerLysAlaLeuGlyLysGluGluGlyLysLysLysLysLysLysLysLys 182  
Db 94757 ATCATCGCAAGTCATGGGAAAGAGAGATGAGATGAGAAACCGCTGGCAGAGACCGT 94698  
QY 183 LysLysIleGluGluTyrLysLysLysLysLysLysLysLysLysLysLysLysPro 202  
Db 94697 GAACGCATGAAGGCTATGCCAGCCAG-----TTGCCACAAGCGCTCAGTGTGTTT 94644  
QY 203 AlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe 222





PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
PI Yamamoto RT, Xu HH;  
DR WPI; 2001-611495/70.  
DR P-PSDB; AAU34876.  
XX  
XX New polynucleotides for the identification and development of  
PT antibiotics, comprise sequences of antisense nucleic acids.  
XX  
XX Claim 27; SEQ ID NO 6372; 511pp; English.  
XX  
XX The invention relates to antisense inhibitors of genes essential to  
CC prokaryotic cellular proliferation, their use in identifying the genes,  
CC their use in the discovery of novel antibiotics, the essential genes  
CC themselves and the encoded proteins. The prokaryotes used are Escherichia  
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,  
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also  
CC useful for the identification of potential new targets for antibiotic  
CC development. The antisense nucleic acids can also be used to identify  
CC proteins used in proliferation, to express these proteins, and to obtain  
CC antibodies capable of binding to the expressed proteins. The proteins can  
CC be used to screen compounds in rational drug discovery programmes. The  
CC antisense nucleic acid sequence is also useful to screen for homologous  
CC nucleic acids which are required for cell proliferation in a wide variety  
CC of organisms. The present sequence encodes an essential prokaryotic  
CC cellular proliferation protein. Note: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 909 BP; 200 A; 262 C; 261 G; 186 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 5.01e-35 Length: 909  
Score: 460.00 Matches: 100  
Percent Similarity: 57.45% Conservative: 62  
Best Local Similarity: 35.46% Mismatches: 104  
Query Match: 27.22% Indels: 16  
DB: 4 Gaps: 4

US-10-724-972A-6352 (1-335) x AAS52735 (1-909)

QY 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64  
DB 76 GTTCAGGACGAAACACGCGACGTTTACATCGAATAAAGCCACACGCGATGCTGTGCTG 135  
QY 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84  
DB 136 GAACCTCTGTCGCGATGCGTGGCGCGGAGCGTCCATCCGATCGTATTGCCGAC 195  
QY 85 AspAsnLysLysAsnArgIleLysProLysArgAspLysIleGlyLysThrSer 104  
DB 196 GATAACGATGCAAAACGATCCTGCGCGAAGTGGTGGCGACCTGAAACCGTGGCAGTCC 255  
QY 105 ValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLysLysProAspLeu 124  
DB 256 GTCCGAACCGCGCGAGCGAGCTGGAGCGCAATTCGCGCTGNAACGACGCTGATC 315  
QY 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLysAsnLysIleAlaProThr 144  
DB 316 ATTGCCGACGAGCAGTCCGATCGCGGGGTTTACATCGCTTGCAGCAATCGCGCGGTA 375  
QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164  
DB 376 CTGCTGCTTAAGTCCCGCAACGAAACCTACGCTGAAATTTGCAATCTCGCGGTATCATC 435  
QY 165 SerLysAlaLeuGlyLysGluGluGlyLysLysLysArgLysLysLysLysLys 184  
DB 436 GCGCAATGCTGGTAAAGACGAGAGATCGAGCGACGCTTGGAAACACATAAGAGAGG 495  
QY 185 IleGluGlyTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204  
DB 496 ATGGCGCAGTGGCCAGCCAG-----CTTCCCAAGGGACACGCGTGGCGCTTTGGCACA 549

QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224  
DB 550 TCACGCGAACAGCAATTCACCTGATCTAGGAGACCTGGACCGGACGCTGCTGGCC 609  
QY 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu 244  
DB 610 TCTCTGGGCTG-----AACGTTCCCGCTGCGATG 639  
QY 245 LysGlyProTyrIleu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263  
DB 640 GCGGTCGTCATGCCATGCCATGCCATGCCATGCCATGCCATGCCATGCCATGCCATGCC 699  
QY 264 MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLys 283  
DB 700 CTCTGCTGTTGCCCACTATCGC-----GAAGAGAGCATTTGTTAAACCTCGACAA 750  
QY 284 AspProValTrpLysLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuArg 303  
DB 751 GATCCGCTCGGCGATGTTAAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 810  
QY 304 AspleuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuVal 323  
DB 811 AACACCTGGCGCGGATGCGCGGTATTTGCTGCGAGCGGTATTCGCGGTGACACGTA 870  
QY 324 GluLeu 325  
DB 871 AAAATC 876  
RESULT 26  
ACA32751  
ID ACA32751 standard; DNA; 909 BP.  
XX  
AC ACA32751;  
XX  
DT 19-JUN-2003 (first entry)  
XX  
DE Prokaryotic essential gene #14408.  
XX  
KW Antisense; ds; prokaryotic essential gene; cell proliferation;  
KW drug design; gene.  
XX  
OS Escherichia coli.  
XX  
PN WO200277183-A2.  
XX  
PD 03-OCT-2002.  
XX  
PF 21-MAR-2002; 2002WO-US009107.  
XX  
PR 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX  
XX WPI; 2003-029926/02.  
DR P-PSDB; ABU28881.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
XX Claim 14; SEQ ID NO 20621; 1766pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC

(1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 909 BP; 200 A; 262 C; 261 G; 186 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	5,01e-35	Length:	909
Score:	460.00	Matches:	100
Percent Similarity:	57.45%	Conservative:	62
Best Local Similarity:	35.46%	Mismatches:	104
Query Match:	27.22%	Indels:	16
DB:	8	Gaps:	4

US-10-724-972A-6352 (1-335) x ACA32751 (1-909)

QY	45	IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu	64
DB	76	GTTCAGGACGACACGGCAGCTTACTCGAATAAAGCCACACGGATGCTGGTGTG	135
QY	65	GluTyrSerPheValAspAlaLeuAspValLysProValLysProValGlyLeuAlaSer	84
DB	136	GAACTCTGTTCCGCGATGCGCTGGCGCGGTGGACGTCATCCCGATCGTATTGCGGAC	195
QY	85	AspAsnLysLysAsnArgIleLysProLysArgAspLysIleGlyLysTyrThrSer	104
DB	196	GATACGATGCAAAACGATCTCGCCGAGTGGCTGGCCACCTGAACCGTGGCAGTCC	255
QY	105	ValGlyThrArgLysGlnProAsnLeuGluGluLysSerLysLeuLysProAspLeuLeu	124
DB	256	GTCGGAACGCGCGCAGCGCAGCTGGAAGCATTGCGGCTGGAACACGACCTGATC	315
QY	125	IleAlaAspAsnAsnArgHisLysGlyLeuLysAspLeuAsnLysIleAlaProThr	144
DB	316	ATTGCGGACAGCATGTCGCGGCTTTACATCGCCTTGGCAGCAATCGCGCGGTA	375
QY	145	IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle	164
DB	376	CTGCTGCTTAAGTCCCGCAACGAAACCTACCTGAAATTTGCAATTCGGCGGTATCATC	435
QY	165	SerLysAlaLeuGlyLysGluGluGlyLysArgLeuGluGluHisAspLysLys	184
DB	436	GGCGAAATGTTGGTGAAGAGGAGATGATGAGGACGCTGCTGGAACAACATAAGAGAGG	495
QY	185	IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal	204
DB	496	ATGCGCGAGTGGGCCAGCCAG-----CTTCCCAAGGGACACGCGTGGCTTTGGCACA	549

QY	205	AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer	224
DB	550	TCACGGGACAGCAATTCACCTGATCTCAGGAGACCTGGACCGCGCTGCTGGCC	609
QY	225	GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu	244
DB	610	TCTCTGGGGCTG-----AACGTTCCCGTGGCGATG	639
QY	245	LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg	263
DB	640	GGGGTGGCTGCAATGCGCTGGCACTGCTGGCGGTCAATCTGCTGGCTGG	699
QY	264	MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLys	283
DB	700	CTGCTGGTTGGCCACTATCGC-----GAAGAGAGCATTTGTTAAACGCTGGCAACA	750
QY	284	AspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg	303
DB	751	GATCCGCTCGCAGATGTTAACCGCGCGCAGACGACGAGTTCCTCGGTCGACAGT	810
QY	304	AspLeuTrpAlaArgSerArgLysLeuIleSerSerGluMetAlaLysGluLeuVal	323
DB	811	AACACCTGGCGCGGATGCGCGTATTTTGTCTGCAGAGCGTATTGCGCGCTGACACGGA	870
QY	324	GluLeu 325	
DB	871	AAAAATC 876	

RESULT 27

AAS46238/c

ID AAS46238 standard; DNA; 10244 BP.

AC AAS46238;

DT 18-DEC-2001 (first entry)

DE DNA encoding novel mar regulated protein (NIMR) #7.

KW mar regulated polypeptide; NIMR; microbial infection; antibacterial; ds.

OS Escherichia coli.

PN WO200170776-A2.

PD 27-SEP-2001.

PF 08-MAR-2001; 2001WO-US007478.

PR 10-MAR-2000; 2000US-0188362P.

PA (TUFT ) TUFTS COLLEGE.

PI Levy SB, Barbosa TM, Alekshun MN;

DR WPI; 2001-602769/68.

XX P-PSDB; AAU29336.

Identifying compounds that modulate a newly identified mar regulated polypeptide activity, useful as antimicrobial compounds, involves contacting the polypeptide with a test compound.

PS Disclosure; Page 178-184; 526pp; English.

The invention relates to a method of identifying compounds that modulate a newly identified mar regulated (NIMR) polypeptide activity. The method comprises contacting an NIMR polypeptide with a test compound under interaction conditions, determining the ability of the compound to modulate the activity or expression of the polypeptide, and selecting the modulators. NIMR nucleic acids and polypeptides are used in the treatment of microbial infections, and in screening for modulators of NIMR expression and activity. These modulators can be used to reduce the infectivity of a microbe on a surface, and the virulence of a microbe in a subject suffering from an infection. AAS46232-AAS46278 represent

```
CC Escherichia coli NMR coding sequences of the invention
SQ Sequence 10244 BP; 2347 A; 2737 C; 2791 G; 2369 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.03e-33 Length: 10244
Score: 460.00 Matches: 100
Percent Similarity: 57.45% Conservative: 62
Best Local Similarity: 35.46% Mismatches: 104
Query Match: 27.22% Indels: 16
DB: 4 Gaps: 4

US-10-724-972a-6352 (1-335) x AAS46238 (1-10244)
QY 45 IleyHisGluGluGluThrLysValProLysHisProLysArgValValValLeu 64
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7735 GTTCAGGACGACGCGCGGTTTACACTCGAAACACGCGCATGTTGGTGGCTG 7676

QY 65 GluTyrSerPheValAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7675 GAACCTCTGTTCCGCGATGCGCTGGCGCGGTGACGTCATCCGATCGGTATTCGCGAC 7616

QY 85 AspAsnLysLysAsnArgIleLysProLeuArgAspLysIleGlyLysTyrThrSer 104
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7615 GATACGATGCAAAACGATCTCTCCCGAAGTGGTGGCACCTGAAACCGTGGCAGTCC 7556

QY 105 ValGlyThrArgLysGlnProAsnLeuGluGluSerLysLysLeuLysProAspLeuIle 124
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7496 GTCCGACGCGCGCGACGCGGCTGGAAGCCATTGCGCTCTGAAACACGACCTGATC 7436

QY 125 IleAlaAspAsnArgHisGlyLysLysLysLysLysLysLysLysLysLysLysLys 144
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7495 ATTGCGGACGACGATGCGCGGTTTACATCGCCTTGGACCAATCGCGCGGTA 7436

QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7435 CTGCTGCTTAAGTCCCGACAAACCTACGCTGAATTTGCAATCTCGCGTATCATC 7376

QY 165 SerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAspLysLys 184
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7375 GCGGAATGTTGGTAAAGGACGAGATGCGAGGACGCTCTGGAACAACATAAGAGG 7316

QY 185 IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7315 ATGGCGGACGTGGCGGACGAG-----CTCCCAAGGACACGCGTGGCTTGGCACA 7262

QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7261 TCACGCGAAACAGCAATTCACCTGATCTACGAGACCTGACCGGACGCGTGGTGGCC 7202

QY 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7201 TCTCTGGGCGTG-----AACGTTCCCGTGGCATG 7172

QY 245 LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7171 GCGGTCGCTGCCATGCGCTCCATCGCGCGGACCACTGCTGGCGTCAATCTGCTGG 7112

QY 264 MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGlyLys 283
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7111 CTGCTGGTTCGCCACCTATCCG-----GAAGAGAGCATTTTAAACGCTGGCAACAA 7061

QY 284 AspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg 303
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7060 GATCCGCTCTGGCAGATGTTAAACCGCGCGGACGAGCAGGTTGCTCGGTCGACAGT 7001

QY 304 AspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuVal 323
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7000 AACACCTGGCGGAGTGGCGGATTTTGTCTGACAGCGTATTGCGCGTGCACGCGTA 6941

QY 324 GluLeu 325
   : : : : :
Db 6940 AAAATC 6935
```

```
RESULT 28
ACA45366
ID ACA45366 standard; DNA; 891 BP.
XX
AC ACA45366;
XX
XX 19-JUN-2003 (first entry)
XX
XX Prokaryotic essential gene #27023.
DE
XX
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX
XX Pseudomonas syringae.
XX
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US0009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX
XX 06-SEP-2001; 2001US-00948993.
XX
XX 25-OCT-2001; 2001US-0342923P.
XX
XX 08-FEB-2002; 2002US-00072851.
XX
XX 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI: 2003-029926/02.
XX
XX P-PSDB; ABU41496.
XX
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 33236; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than S. aureus, S. typhimurium,
XX K. pneumoniae or P. aeruginosa. The present sequence is one of the target
XX prokaryotic essential genes. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
```



Score: 405.00 Matches: 113  
Percent Similarity: 51.74% Conservative: 65  
Best Local Similarity: 32.85% Mismatches: 132  
Query Match: 23.96% Indels: 34  
DB: 8 Gaps: 12

US-10-724-972A-6352 (1-335) x ABT15032 (1-990)

QY	8	LeuLysLeuSerValleGlyLeuLeuPheValLeuLeuAlaThrAlaAlaCysGly	27
DB	13	ATTAAATGCTT--GTTGTTACGCTTGTCTTCTACTTGT--TTAGCAGGATGTAGT	66
QY	28	AsnAsnSerSerSerAsnSerSerLysGluSerSerLysAspGlyValGluLeuLysHis	47
DB	67	GGGAATTCAAATAAACAATCATCTGATACAAAGATAAGAAACAACATTCATTAACAT	126
QY	48	GluGluGlyThrThrLysValProLysHisProLysArgValValValLeuGluTyrSer	67
DB	127	GCAATGGGTACACTGAAATTAAGGAAACCAAGCGTGTGTGTACGCTATATCAAGGT	186
QY	68	PheValAspAlaLeuValAlaLeuAspValLysProValGlyLeuAlaAspAsnLys	87
DB	187	GCCACTGACCTGCTGTATCTTTAGGTGTAAACCTGTAGGTGCTGTAGTATCATGGACA	246
QY	88	LysAsnArgilelleLysProLysArgPheLysLysGlyLysThrSerValGlyThr	107
DB	247	CAAAAACCGAAATTCGATACATATAAATAATGATTTAAAGATACTAAGATTGTAGTCAA	306
QY	108	ArgLysGlnProAsnLeuGluLysSerLysLeuLysProAspLeuLeuLeuAlaAsp	127
DB	307	GAACCTGCACCTTAATTAGAGGAATCTCTAATTAAACCGGACTTAATTGTCGCTCA	366
QY	128	AsnAsnArgHisLysGlyLysLysArgLeuLeuLysLeuLysLeuLysLeuLysLeu	147
DB	367	AAAGTTAGAATGAAAAGAGCTGAAGATTTACTTTAAAGATGACGATGATAAGTGC	420
QY	148	LysSerPheAspGlyAspTyrAsnGluAsnLeuAspAlaPheLysThrLysLeuAla	167
DB	421	---TCTACTGATACAGTTTCAAAATTCAAA--GATACAACATAAGTTAATGGGAAGCT	474
QY	168	LeuGlyLysGluGluGluLysLysArgLeuGluGluHisAspLysLysLeuGluLys	187
DB	475	TTAGGAAAGAAAAGAGCTGAAGATTTACTTTAAAGATGACGATGATAAGTGC	534
QY	188	TyrLysLysGluLysLeuMetAspLysAsnGlnLysValLeuProAlaValAlaLys	207
DB	535	TTCAAAA-----GATGAAAAGCAAAAGTATAAGATGATGCGCATTTGAAA	582
QY	208	SerGlyLeuLeuAlaHisProSerAsn-----SerTyrValGly	220
DB	593	GCTTCAGTTGTTAACTTCGCTGCTGATCATACAAGAAATTTATGCTGCTGATATGCTGT	642
QY	221	GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu	240
DB	643	GAATCTTAAATGATTTAGGATTTCAACGT-----AATAAGACTTA	684
QY	241	SerLysTyrLeuLysGly-----ProTyrLeuGlnMetAsnThr-----GluThrLeuSer	257
DB	695	CAAAAACAAAGTTGATAATGTTAAAGATATATATCACTTACATCAAGAAAGCAATTC	744
QY	258	GlnValAsnProGluArgMetPheLeuMetThrAsnLysAlaSerSerAsnGluProSer	277
DB	745	TTAATGAACGCTGATCATATTTTGTAGTAAATATCAGATCCAAATCGAAAGATGCTGCA	804
QY	278	Leu-----LysGluLeuGluLysAspProValTpyLysLysLeuAsnAla	292
DB	805	TTAGTTAAAAAGACTGAAGCAATGGACTTCAAGTAAAGAGTGGAAAAATTTAGACGCA	864
QY	293	ValLysAsnGlnArgVal---AspLeuAspArgAspLeuTyrPheLysArgGly	311
DB	865	GTTAAANAACCAACCAAGTATCTGATTTAGATGAATCACTTGGAACTTAGCTGGCGGA	924
QY	312	LeuLeuSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAspSerLys	331

925 TATAAATCTTCATTAAAACTTATTGACGATTATATGAA-----AAGTTAAATATTGAA 978

332 LysAspAsnLys 335

979 AAACAATCAAAA 990

RESULT 30

AAZ22850

ID AAZ22850 standard; DNA; 999 BP.

XX AC AAZ22850;

XX 20-DEC-1999 (first entry)

DE Staphylococcus aureus cbrA DNA.

XX Infection; detection; diagnosis; screening; antibiotic; resistance;

KW methicillin; MRSA; ds.

XX Staphylococcus aureus.

OS Location/Qualifiers

Key 7..999 a

FT CDS /tag= a

FT /product= "S. aureus cbrA protein"

XX WO9947662-A1.

XX 23-SEP-1999.

XX 18-MAR-1999; 99WO-US006199.

XX 20-MAR-1998; 98US-0078682P.

PR 01-APR-1998; 98US-0080296P.

PR 07-MAY-1998; 98US-0084674P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (LUDW-) LUDWIG INST CANCER RES.

XX Simpson AJG, Choi GH;

WPI; 1999-580304/49.

DR P-PSDB; AA42331.

XX Staphylococcus aureus genes and polypeptides, vectors and methods of recombinant production.

XX Claim 1; Page 24; 111pp; English.

XX This sequence represents Staphylococcus aureus cbrA DNA. The cbrA protein is predicted to have a molecular weight of about 36.8 kD and, along with cbrB (AA42332) and cbrC (AA42333), is thought to be involved in iron regulation, based on amino acid sequence homology with known iron regulator proteins. S. aureus is a ubiquitous pathogen which causes infections in burns, cellulitis, eyelid infections, food poisoning, joint infections, neonatal conjunctivitis, osteomyelitis, skin infections, surgical wound infection, scalded skin syndrome and toxic shock syndrome. S. aureus is increasingly becoming resistant to known antibiotics, with methicillin-resistant strains generally being multiply drug resistant. Methicillin-resistant S. aureus (MRSA) poses serious infection control problems, with many strains being multiresistant against virtually all antibiotics with the exception of the vancomycin-type glycopeptide antibiotics. The protein may be useful to screen potential antagonists which could be used as antibiotics and it may be used as a vaccine to prevent or attenuate an infection caused by a member of the Staphylococcus genus. The protein, or antibodies against it can be used in immunoassays to detect Staphylococcus in a biological sample. Probes and primers derived from the nucleic acid sequences may also be used to detect Staphylococcus nucleic acids in a biological sample

XX Sequence 999 BP; 402 A; 142 C; 180 G; 275 T; 0 U; 0 Other;



Alignment Scores:

Pred. No.:	1.19e-29	Length:	999
Score:	405.00	Matches:	113
Percent Similarity:	51.74%	Conservative:	65
Best Local Similarity:	32.85%	Mismatches:	132
Query Match:	23.96%	Indels:	34
DB:	2	Gaps:	12

US-10-724-972A-6352 (1-335) x AAZ222850 (1-999)

Qy	8	LeuLysIleLeuSerValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCysGly	27
Db	19	ATTAATAATGCTT---GTTGTACGCTTGCTTTCTACTTCTT---TTAGCAGGATGAGT	72
Qy	28	AsnAsnSerSerSerAsnSerSerLysGluSerSerLysAspGlyValGluIleLysHis	47
Db	73	GGGAATTCAAATTAACAACATCATCTGATACAAAGATAAGAAACAACATTCATTAACAT	132
Qy	48	GluGluGlyThrThrLysValProLysHisProLysArgValValValLysLeuLysSer	67
Db	133	GCAATGGGTACAACTGAAATTAAGGGAAACCAAGCGTGTGTTCAGCTATATCAAGT	192
Qy	68	PheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLys	87
Db	193	GCCACTGACGTGCGTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAATCATGGACA	252
Qy	88	LysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyThrSerValGlyThr	107
Db	253	CAAAAACCCAAATTCGAATACATAAAAAATGATTTAAAGATACCTAAGATTGTAGTC	312
Qy	108	ArgLysGlnProAsnLeuGluIleSerLysLeuLysProAspLeuIleAlaAsp	127
Db	313	GAACCTGCACCTAACTTAGAGGAAATCTCTAAATTAAGCCGACTTAATTCGCGTCA	372
Qy	128	AsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeu	147
Db	373	AAAGTTAGAAATGAAAAGTTTACGATCAATTTATCTAAATCGCAACCAACAGTT---	426
Qy	148	LysSerPheAspGlyAspTyAsnGluAsnIleAspAlaPheLysThrIleSerLysAla	167
Db	427	---TCTACTGATACAGTTTTCAAATTCAAA---GATCAACCTAAGTTAATGGGAAAGCT	480
Qy	168	LeuGlyLysGluGluGluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGlu	187
Db	481	TTAGGGAAGAAAAAGAGCTTACTTAAAAAGTACGATGATAAAGTAGCTGCA	540
Qy	188	TyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLys	207
Db	541	TTCCAAAAA---GATGCAAAAGCAAAAGTATAAAGATGCATGCCCATTTGAA	588
Qy	208	SerGlyLeuLeuAlaHisProSerAsn-----SerTyrValGly	220
Db	589	GCTTCAGTTGTAACTTCCTCGCTGATACATAACAAGATTTATGCTGGTATATGCTG	648
Qy	221	GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu	240
Db	649	GAATCTTAATGATTTAGATTCAACGT-----AATAAAGACTTA	690
Qy	241	SerLysTyrLeuLysGly-----ProTyrLeuGlnMetAsnThr---GluThrLeuSer	257
Db	691	CAAAAAACAAGTTGATATGTTAAAGATATTATCCAACTTACATCTAAGAAAGCATTTCCA	750
Qy	258	GlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSer	277
Db	751	TTAATGAACGCTGATCATATTTTGTAGTAAATTCAGATCCAAATCGGAAGATGCTGCA	810
Qy	278	Leu-----LysGluLeuGluLysAspProValTrpLysLysLeuAsnAla	292
Db	811	TTAGTTAAAAAGACTGAAAGCGAATGGCATTCAGTAAAGAGTGGAAAAATTTAGACGCA	870
Qy	293	ValLysAsnGlnArgVal---AspIleLeuAspArgAspLeuTrpAlaArgSerArgGly	311
Db	871	GTTAAAAACAACAGTATCTGATGATTTAGATGAATCACTTGGAACTTAGCTGGCGGA	930

Qy	312	LeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAspSerLys	331
Db	931	TATAAAATCTTCATTAATAAACTTATTGACGATTATATGAA-----AAGTTAAATATTGAA	984
Qy	332	LysAspAsnLys	335
Db	985	AAACAATCAAAA	996

RESULT 31

AAZ19889<sup>x</sup>  
ID AAZ19889 standard; DNA; 999 BP.  
XX  
XX  
AC  
AC AAZ19889;  
XX  
XX  
DT 06-DEC-1999 (first entry)  
XX  
XX  
DE Staphylococcus aureus iron regulation gene cbra.  
XX  
XX Cbra gene; infection; therapy; diagnosis; vaccine; antibiotic;  
KW iron regulation; ss.  
KW

OS Staphylococcus aureus.

AA	Key	Location/Qualifiers
FH	CDS	7..999
FT		/*tag= a
FT		

WO9947639-A2.

XX  
PD  
23-SEP-1999XX  
DE 10-MAY-1968

XX  
XX

PR 01-APR-1998; 98US-0080296P.

XX  
07-111-1336, 3803-00846/4P.

PA (HUMA-) HUMAN GENOME SCI INC.  
PA (LUDW-) LUDWIG INST CANCER RES

XX  
PT  
Simpson A.T.C  
Choi CH.

**XX**

DR P-PSDB; AAY31824.

PT New isolated *Staphylococcus* nucleic acid molecules, used to develop  
PT products for the diagnosis, prevention and treatment of *Staphylococcal*  
PT infections.

PS Claim 1; Page 24; 102pp: English.

This is the nucleotide sequence of *Staphylococcus aureus* strain ISP3 (ATCC 202108) genomic DNA including the novel cbrA gene that codes for a 330-amino acid protein (see AAY31824) of predicted mol. wt. 36.8 kDa. The sequence was obtained from overlapping clones BTAC444 and BTAG354, which span a single operon containing the cbrA, cbrB and cbrC genes (see AAY1989-91). CbrA shows sequence homology to known genes involved in iron regulation. The invention provides 11 novel genes (see AAY1982-92) of *S. aureus* and the polypeptides they encode (see AAY1811-27). Also provided are vectors, host cells, antibodies and hybridomas. The invention further relates to screening methods for identifying agonists and antagonists of *S. aureus* polypeptide activity, and to diagnostic methods for detecting *Staphylococcus* nucleic acids, polypeptides and antibodies in a biological sample. Antagonists of cbrA may be useful as antibiotics to treat infections of *S. aureus* and other *Staphylococcus* spp. Also provided are novel vaccines for the prevention or attenuation of infection by *Staphylococcus*. The isolated nucleic acid molecule is also useful for generating probes and primers, and in the recombinant production of cbrA protein.

Sequence 999 BP; 402 A; 142 C; 180 G; 275 T; 0 U; 0 Other;



Query Match:	23.96%	Indels:	34
DB:	10	Gaps:	12
US-10-724-972A-6352 (1-335) x ADF43555 (1-999)			
QY	8	LeuLysIleLeuSerValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCysGly	27
DB	19	ATTAAATGCTT---GTTCAGCGCTTGCTTTCCCTACTGTT---TTAGCAGGATGTAGT	72
QY	28	AenAsnSerSerSerAsnSerSerLysGluSerSerLysAspGlyValGluIleLysHis	47
DB	73	GGGAATTCATAAACAACATCATCTGATCAACAAGATAAGGAACAACCTTCAATTAACAT	132
QY	48	GluGluThrThrLysValProLysHisProLysArgValValValLeuGluTyroser	67
DB	133	GCAATCGGTACCACTGAATTAACGGGAACCAAGCGGTGTGACCTCATATCAAGGT	192
QY	68	PheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsnLys	87
DB	193	GCCACTGCCTCGCTGATCTTTAGGTGTTAAACCTGTAGGTGTGTAGATCATGACA	252
QY	88	LysAsnArgIleLysProLeuArgAspLysIleGlyLysThrSerValGlyThr	107
DB	253	CAAAAACCGAAATTCGAATCATATAAATAATGATTAAAGATACCTAGGTGAGTCAA	312
QY	108	ArgLysGlnProAsnLeuGluGluLysSerLysLysLysProAspLeuIleAlaAsp	127
DB	313	GNACCTGCACCTAACCTTAGAGAATCTCTAAATTAACCGGACTTAATTTGCGGTCA	372
QY	128	AsnAsnArgHisLysGlyIleTyLysAspLeuAsnLysIleAlaProThrIleGluLeu	147
DB	373	AAAGTTAGAAATGAAAAAGTTTACGATCAATATCTAAATTCGCCAACAGTT-----	426
QY	148	LysSerPheAspGlyAspTyAsnGluAsnIleAspAlaPheLysThrIleSerLysAla	167
DB	427	--TCTACTGATACAGTTTTTCAAAATTCAAA--GATACAACCTAAGTTAATGGGGAAGCT	480
QY	168	LeuGlyLysGluGluGluGlyLysLysArgLeuGluLysAspLysIleGluGlu	187
DB	481	TTAGGGAAGAAAAGAACGCTGAAGATTTACTTAAAGATGACGATGATAAGTAGTGCA	540
QY	188	TyrLysLysGluIleThrMetAspLysGlnLysValLeuProAlaValAlaLys	207
DB	541	TTCCAAAA-----GATGCAAAAGCAAGTATAAGATCGATCGCCATTTGAAA	588
QY	208	SerGlyLeuLeuAlaHisProSerAsn-----SertYrValGly	220
DB	589	GCTTCAGTTGTAACTTCGCTGATCATCAAGAAATTTATCTCGTGGTGAATGCTGT	648
QY	221	GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu	240
DB	649	GAATCTTAATGATTTTAGGATTCACACGT-----AATAAGACTTA	690
QY	241	SerLysTyrLeuLysGly-----ProTYrLeuLysMetAsnThr-----GluThrLysSer	257
DB	691	CAMAAACAGTTGTAATGTGTAAGATATTTATCCAACTTACATCTCAAGAAAGCATTC	750
QY	258	GlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSer	277
DB	751	TTAATGAACCTGATCATATTTTGTAGTAAATCAGATCCAAATCGCAAGATGCTGCA	810
QY	278	Leu-----LysGluLeuGluLysAspProValTrpLysLysLeuAsnAla	292
DB	811	TTAGTTAAAAAGACTGAAAGCGAATGGACTTCAGGTAAGAGTGGAAAAATTTAGACCA	870
QY	293	ValLysAsnGlnArgVal---AspIleLeuAspArgAspLeuTrpAlaArgSerArgGly	311
DB	871	GTTAAAAACAACCAAGTATCTGATGATTTAGATGAATCACCTTGGAACTTAGCTGCCGA	930
QY	312	LeuIleSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAspSerLys	331
DB	931	TATAAATCTTCATTAACACTTATTGACGATTTTATATGAA-----AAGTTAAATATGAA	984



Best Local Similarity: 32.85% Mismatches: 132  
Query Match: 23.96% Indels: 34  
DB: 2 Gaps: 12

US-10-724-972a-6352 (1-335) x AAV74549 (1-3775)

QY	8	LeuLysIleLeuSerValleGlyLeuLeuPheValLeuIleAlaThrAlaCysGly	27
DB	61	ATTAAATGCTT---GTTGTAGCTTGTCTTCTACTTGT---TTAGCAGGATGAGT	114
QY	28	AsnAsnSerSerSerAsnSerLysGluSerSerLysAspGlyValGluIleLysHis	47
DB	115	GGGATTCATTAACAAATCATCTGATAACAAAGATAAGGAACAACATTCATTAACAT	174
QY	48	GluGluGlyThrThrLysValProLysHisProLysArgValValLeuGluTyrSer	67
DB	175	GCAATGGGTACAACTGAATTAAGCGAACAACCAACGCTGTGTAGCTATATCAAGT	234
QY	68	PheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsnLys	87
DB	235	GCCACTGACGCTGCTGTATCTTAGGTGTAAACCTGTAGGTGTAGTAATCATGGACA	294
QY	88	LysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThr	107
DB	295	CAAAAACCGAATTCGATACATAAAAATGATTTAAAGATACATAAGATTGTAGTCAA	354
QY	108	ArgLysGlnProAsnLeuGluIleSerLysLeuLysProAspLeuIleIleAlaAsp	127
DB	355	GAACCTGCACCTTAAGAGAAATCTTAATTAACACCGACTTAATGTGCGGTCA	414
QY	128	AsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeu	147
DB	415	AAAGTTGAAATGAAAAGTTTACGATCAATTTATTAATTCGACCAACAGTT-----	468
QY	148	LysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAla	167
DB	469	---TCTACTGATACAGTTTTCAAATTCAAA---GATACAACTAAGTTAATGGGGAAGCT	522
QY	168	LeuGlyLysGluGluGlyLysArgLeuGluGluHisAspLysLysIleGluGlu	187
DB	523	TTAGGGAAGAAAAGAGCTGAAGATTACTTAAAGATGATGATGAATAGTACTGCA	582
QY	188	TyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLys	207
DB	583	TTCCAAAA-----GATCAAAAGCAAGTATAAGATGCATGCCATTGAA	630
QY	208	SerGlyLeuLeuAlaHisProSerAsn-----SerTyrValGly	220
DB	631	GCTTCAGTTGTAACCTCCGCTGCTGATCATACAGATTTTATGCTGTGATATGCTGCT	690
QY	221	GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu	240
DB	691	GAATCTTAATGATTAGGATTCAACGT-----AATAAAGACTTA	732
QY	241	SerLysTyrLeuLysGly-----ProTyrLeuGlnMetAsnThr---GluThrLeuSer	257
DB	733	CAAAAACCAAGTTGATATGTAAGATATTTATCAACATTTATGCTGTGATATGCTGCT	792
QY	258	GlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSer	277
DB	793	TTAATGAACGCTGATCATATTTTGTAGTAAATCAGATCCAAATGCCAAGATGCTGCA	852
QY	278	Leu-----LysGluLeuGluLysAspProValTyrLysLeuAsnAla	292
DB	853	TTAGTTAAAGACTGAAGCAATGAGCTCAAGTAAAGAGTGGAAAATTTTAGCGCA	912
QY	293	ValLysAsnGlnArgVal---AspIleLeuAspArgAspLeuTyrPalaArgSerArgGly	311
DB	913	GTTAAAAACCAACCAATGATCTGATGATTAGTGAATCACTTGAACCTAGCTGGCGGA	972
QY	312	LeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAspSerLys	331
DB	973	TATAAATCTTCTTAAACACTTATTGACGATTTATATGAA-----AAGTTAAATATTGAA	1056

QY 332 LysAspAsnLys 335  
DB 1027 AACAAATCAAAA 1038

RESULT 35

ACA21545  
ID ACA21545 standard; DNA; 972 BP.  
XX  
AC ACA21545;  
XX  
DT 19-JUN-2003 (first entry)  
XX  
DE Prokaryotic essential gene #3202.  
XX  
KW Antisense; ds; prokaryotic essential gene; cell proliferation;  
KW drug design; gene.  
XX  
OS Bacillus anthracis.  
XX  
PN WO200277183-A2.  
XX  
PD 03-OCT-2002.  
XX  
PF 21-MAR-2002; 2002WO-US009107.  
XX  
PR 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX  
PA (ELIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
WPI; 2003-029926/02.  
P-PSDB; ABU17675.

New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

Claim 14; SEQ ID NO 9415; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target

CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SO Sequence 972 Bp; 405 A; 147 C; 183 G; 237 T; 0 U; 0 Other.

Alignment Scores:	2.6e-28	Length:	972
Pred. NO.:	391.00	Matches:	116
Score:	51.93%	Conservative:	59
Percent Similarity:	51.93%	Mismatches:	122
Best Local Similarity:	34.42%	Indels:	13
Query Match:	28.14%	Gaps:	10
DB:	8		

UUS-10-724-972A-6352 (1-335) x ACA21545 (1-972)

QY	9	LysIleLeuSerValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCysGlyAsn	28
DB	7	AAAATTTCTCAGTATT-----TTTCATAGTAGITTTTCTATTTCGCTGTTGGATGC	54
QY	29	AsnSerSerSerAsnSerSerLysGluSerSerLysAsp-----GlyValGlu	44
DB	55	GGACAGCAAAAGAGAGGAGAAAAGAAACAAACCGGACCAATAAAATCAAGCTATAACA	114
QY	45	IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu	64
DB	115	ATTAAACACGCTGAAAGGGGAAACGAAGTTAGATAAACCCAGCGAAAAAAGTAGTTGACTT	174
QY	65	GluTyrSerPheValAspAlaLeuValAlaIleuAspValLysProValGlyIleAlaAsp	84
DB	175	GAATGGGTATATTTCAGAAAGACTTTATTAGCACTTTGGTGTTCAGCCAGTAGGAGTAGGCAGAC	234
QY	85	AspAsnLysLysAsnArgIleIle-----LysProLeuArgAspLysIleGly	100
DB	235	ATTAAGAATTTATAATAAATGGGTAAATACAAAAACAAACCGAGTAAAGATGTTGTA--	291
QY	101	LysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluIleSerLysLeuLys	120
DB	292	-----GATGTCGGGACACGTCACCAACCAAACTTAGAAGAAATTAGCCGTTTAAAA	342
QY	121	ProAspLeuIleIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLys	140
DB	343	CCAGATTAAATATATACACAGCTTCATTCGCTGGTAAAGCAATATAAATGAATTAGACAA	402
QY	141	IleAlaProThrIleGluLeu-----LysSerPheAspGlyAspTyrAsnGluAsn	157
DB	403	ATTGCACCAACAGTTATGTTTGATCCATCAACCAAGCAATAACGATCACTTTGCTGAAATG	462
QY	158	IleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGluGlyLysArg	177
DB	463	ACAGAAACATTTAAACAAATTTGCAAAAGCAGTTGGAAAAGAAAGAAAGAGGTAAAAAGTA	522
QY	178	LeuGluGluHisAspLysLysIleGluIuTyrLysLysGluIle-----Thr	193
DB	523	TTAGCTGATATGGATAAAGCATTCCTCGATGCCAAAGCAAAAATTTGAAAAAGCAGACTTA	582
QY	194	MetAspLysAsnGlnLysValLeuProAlaValAlaLysSer-----Gly	209
DB	583	AAAGATAAAAAATCATCGCAATGCGCAAGCATTTTACTGCTAAAAAATGTGCCACATTCGT	642
QY	210	LeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLys	229
DB	643	ATCTTAAC-----GACAAATTTCTTTAGCTTTTACAGTTACAAAAAATTAGGT-----	690
QY	230	GluAlaLeuSerAspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeu	249
DB	691	-----TTAAACAATACTTTTGAAGCAGGAAAAATCTGAG-----CCTGATGGTTTCAA	738
QY	250	GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg---MetPheIleMetThr	268
DB	739	CAAAACAACTGTGGAATTCATTACAAAGTGTCAAGATTCAACTCAAACTTCATTATGTTAGGC	798

SQ Sequence 944 BP; 332 A; 197 C; 207 G; 208 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 2.6e-27 Length: 944  
Score: 380.50 Matches: 109  
Percent Similarity: 52.92% Conservatively: 63  
Best Local Similarity: 33.54% Mismatches: 128  
Query Match: 22.51% Indels: 25  
DB: 6 Gaps: 13

US-10-724-972A-6352 (1-335) x ABK74770 (1-944)

```
QY      8 LeuLysIleLeuSer---ValIleGlyLeuLeuPheValLeuIleAlaThrAlaAcyS 26
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1 TTGAAATAACTATCTTTATGATTATGGCATTGTAACCGTTCTTCGTAGCGCGCTGC 60
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY     27 GlyAsnSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 45
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      61 GGAACAAAGCTGCTGACACAGATGCAAAATCTTCTCTTAAATAAGAAACCGTCACAAT 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY     46 LysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeuGlu 65
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     121 AAACACACAGCGACACACAGATGACGAAACCCGAAACCCGAAACAGTCGTCTTCAT 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY     66 TyrSerPheValAspAlaLeu-----ValAlaLeuAspValLysProValGlyIleAla 83
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     181 TTGGAATCTTGACACGCTTGATGCTTGGCTTAAGCGACCGCTTTCGCGCTTCG 240
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY     84 AspAspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThr 103
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     241 AAACAAACCTTCGCTCT---TACTTGAAAAATATAAGAC-----AGCAAAATTTGAA 291
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY    104 SerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLysLeuLysProAspLeu 123
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     292 AGCGTCGGCGGACTAAAGAGCCAGATTTTGAATAAATGCTGAGCTTGAGCTGATTTG 351
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY    124 IleIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaPro 143
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     352 ATCATCATCGAA---GGCAGACAGCTCTGAACAGATATGACGAGTTTAAAGAGATCGCTCG 408
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY    144 ThrIleGluLeu-----LysSerPheAspGlyAspTyrAsnGluAsnIleAsp 159
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     409 ACGATCGATGACCTTGACGAGAGACATACATGACAGTTTAAAGAAATAACCGAA 468
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY    160 AlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGluLysLysArgLeuGlu 179
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     469 -----AACATTGGAAAAATCTTCGCAAGAGAGATGCTGTCNAAGAGCCCTTCGC 519
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY    180 GluHisAspLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLys 199
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     520 AAAATCGATGAAAGTAGAGAAAGCCCAACAG---CTGGCGAAGAAAAAGCGGCAAA 576
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY    200 ValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrVal 219
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     577 GGATTGGTCTGCTTAACGAGCGATGAAGCAATACAGCGGTATGGAAGCGGCTCAAGATTC 636
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY    220 GlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGly 239
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     637 GGCTTGATCCAGATGCTCGCGCTT-----GAACCTGTCGATAAAAC 681
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY    240 LeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnVal 259
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     682 ATCGAGGTTTCCCTTCACGGCCAA-----AGCGTCTTCCAGATATATCTCTGAAAG 735
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY    260 AsnProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluPro---SerLeu 278
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     736 AATCCTGATTACCTGTTGTCATCGACCGAGCGCTGCCATCGCGGAAAAAGGAACAGCA 795
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY    279 LysGluLeuGluLysAspProValTyrPheLysLysLeuAsnAlaValLysAsnGlnArgVal 298
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     796 AAGAAGTCTATCGAAATGAATGTAATCAACAAGAGCGCTATATAAAGCGGCATATC 855
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY    299 AspIleLeuAspArgAspLeuTrpAlaArgSer---ArgGlyLeuIleSerSerGluGlu 317
```

```
Db      856 ACATACCTTGATCCAGGCTACTGGTACCTGTCCAGCGGAGGCGCTTGAATCCGTATCTGAA 915
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY     318 MetAlaLysGluLeu 322
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     916 ATGGTGAAGAAGTC 930
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 37
AAS71379/c
ID AAS71379 standard; cDNA; 2209 BP.
XX
XX AAS71379;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #7183.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Dmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG07192.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 7183; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2209 BP; 480 A; 603 C; 587 G; 539 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.56e-26 Length: 2209
Score: 375.00 Matches: 96
Percent Similarity: 50.17% Conservatively: 52
```

XX	DNA encoding novel human diagnostic protein #8842.
DE	Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX	food supplement; medical imaging; diagnostic; genetic disorder; ss.
KW	Homo sapiens.
XX	WO200175067-A2.
XX	11-OCT-2001.
PD	30-MAR-2001; 2001WO-US008631.
XX	31-MAR-2000; 2000US-00540217.
XX	23-AUG-2000; 2000US-00649167.
PR	(HYSE-) HYSEQ INC.
XX	Drmanac RT, Liu C, Tang YT;
PI	WPI; 2001-639362/73.
XX	P-PSDB; ABG08851.
DR	New isolated polynucleotide and encoded polypeptides, useful in
DR	diagnostics, forensics, gene mapping, identification of mutations
XX	responsible for genetic disorders or other traits and to assess
XX	biodiversity.
PS	Claim 1; SEQ ID NO 8842; 103pp; English.
XX	The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC	sequences. (I) is useful as hybridisation probes, polymerase chain
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC	and in recombinant production of (II). The polynucleotides are also used
CC	in diagnostics as expressed sequence tags for identifying expressed
CC	genes. (I) is useful in gene therapy techniques to restore normal
CC	activity of (II) or to treat disease states involving (II). (II) is
CC	useful for generating antibodies against it, detecting or quantitating a
CC	polypeptide in tissue, as molecular weight markers and as a food
CC	supplement. (II) and its binding partners are useful in medical imaging
CC	of sites expressing (II). (I) and (II) are useful for treating disorders
CC	involving aberrant protein expression or biological activity. The
CC	polypeptide and polynucleotide sequences have application of mutations
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC	coding sequences of the invention. Note: The sequence data for this
CC	patent did not appear in the printed specification, but was obtained in
CC	electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	Sequence 4392 BP; 1063 A; 1157 C; 1181 G; 990 T; 0 U; 1 Other;
SQ	
Alignment Scores:	
Pred. No.:	6.05e-26 Length: 4392
Score:	375.00 Matches: 96
Percent Similarity:	50.17% Conservative: 52
Best Local Similarity:	32.54% Mismatches: 115
Query Match:	22.19% Indels: 32
DB:	Gaps: 5
US-10-724-972A-6352 (1-335) x AAS73038 (1-4392)	
QY	45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64
DB	70 GTTCAGGACGAACACGCGACGTTACACTCGAATAAACCCACACGAGATTGTGGTCTG 129
QY	65 GluTySerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84
DB	130 GAATCTCTGTTGCCGATCGCTGGCGCGCGTGCATCCGATCGGTATTGCCGAC 189
RESULT 38	
AAS73038	
ID	AAS73038 standard; cDNA; 4392 BP.
XX	
AC	AAS73038;
XX	
DT	13-FEB-2002 (first entry)



QY	85	AspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSer	104
DB	190	GATAACGATGCAAAACGCATCTCTCCCGAAGTGGTGGCGACCTGAAACCGTGGCAGTCC	249
QY	105	ValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIle	124
DB	250	GTCGGAACCGCGCGCAGCGTGGAAAGCCATTGCCGCTCTGAAACACACCTGATC	309
QY	125	IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr	144
DB	310	ATTGCCGACAGTCGCATCGCGGGGTTCATATCGCTTCGACAAATCGCGCGGTA	369
QY	145	IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle	164
DB	370	CTGCTGCTTAACTCCGCAACAAACCTACGCTGAAATTTTGCATCTCGCGCTATCATC	429
QY	165	SerLysAlaLeuGlyLysGluGluGlyLysArgLeuGluGluHisAspLysLys	184
DB	430	GGCGAAATGGTGGGTAAAGCCAGAGATGCAGGCACGCTGTGAAACAACATAAGAGAGG	489
QY	185	IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal	204
DB	490	ATGGCGCAGTGGCGCAGCCAG-----CTTCCAAAGGACACCGCTGGCTTGGCAC	543
QY	205	AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer	224
DB	544	TCAGCGAAGACAGCAATTCACACTGCATACTCAGGAGACCTGGACCGCGAGCGTGTGGCC	603
QY	225	GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu	244
DB	604	TCTCTGGGGCTG-----AACGTTCCCGCTGCGCATG	633
QY	245	LysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMet	264
DB	634	CGCGGTGGCTCATCGCGCTCTCGCTGGTGGCGGCTTCATGTCACCGCGATGT	693
QY	265	PheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGluLysAsp	284
DB	694	TCTATCCATGCAATTACGGTTTACATCAACACACCCCTGCTC-----	735
QY	285	ProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeu-----	301
DB	736	-----TGGACCGCGGGTTCTGTGATCTGTTGGCTGGTGGTGTCTGAAATGACC	789
QY	302	-----AspArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGlu	316
DB	790	GACGAAGCCGCTGAAGATGCGAAACTGGTTGGGTTCCGCACACACAGCTGACGCAAGAA	849
QY	317	-----GluMetAlaLysGluLeuValGlnLeuSerLys	327
DB	850	TACGATCACATTAAAGACGTTAAACAGATCTGCTGAATCTGCTGAATGCTGAAA	894

RESULT	39
AAV74666	
ID	AAV74666 standard; DNA; 668 BP.
XX	
XX	AAV74666;
XX	
DT	16-MAR-1999 (first entry)
DT	
XX	
DE	Staphylococcus aureus contig SEQ ID #355.
XX	
KW	Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW	cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW	skin infection; surgical wound infection; scalded skin syndrome;
KW	toxic shock syndrome; ds.
XX	
OS	Staphylococcus aureus.
XX	
XX	
FH	Key Location/Qualifiers
FT	misc_feature 61..120
FT	/tag= a
FT	/note= "these bases represent a line of missing text in
FT	

**FT FT FT XX PN XX PD XX PF PR XX PA XX PI XX DR PT PT PT XX PS XX CC CC CC CC CC CC CC CC CC CC CC CC CC CC CC CC CC CC**

Alignment Scores:  
Pred. No.:  
Score:  
Percent Similarity:  
Best Local Similarity:  
Query Match:  
DB:

US-10-724-972A-6352 (1-335) x AAV74666 (1-668)

210	LeuLeuAlaHisProSerAenSerTyrValGlyGlnPheLeuSerClnLeuGlyPheIys	229
QY		
Db	2 TTATTAGCACATCCAAACTATTCATATGTGGCAATTTTAAACGCACTAGGATTTTAA	61
QY	230 GluAlaLeuSerAspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeu	249
Db	62 NNN	121
QY	250 GlnMetAsnThrGluThrLeuSerGlnValAsnProGluAtrGlyMetPheIleMetThrAsn	269
Db	122 CAATTAGACACTGACATTTAGCTGATTTAAATCCAGACCGTGATGATCATTTATGACAGAT	181
QY	270 LysAlaSerSerAsnGluProSerLeuLysGluLeuGluLysAspProValTrpLysLys	289
Db	182 CATGCTAAAAAGAGTTCTGCTGAATTCACAGAAGTTTACAGAAGAGATGCAACATCGAAAAAG	241
QY	290 LeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSer	309

the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

EP786519-A2.

30-JUL-1997

07-JAN-1997: 97EP-00100117

05-JAN-1996: 96UIS-00009861P

(HTMA-) HITMAN GENOME SCT TNC

Kingsch CA  
Choi CH  
Barnes  
Gardner  
Dillon  
Meyer

WPT: 1007 374022/25

Polyminalceid (a)

stored on computer readable medium and used in the production of anti-

100

[illegible]

The DNA sequences are recorded on a computer readable medium, preferably selected from a floppy or hard disk, random access memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using the *S. aureus* DNA sequences allows putative functions to be assigned so that the protein-encoding or regulatory regions of commercial, therapeutic or industrial importance can be obtained. Specifically, sequences which are likely to encode antigens have been identified and these polypeptides can be used in a vaccine composition against *S. aureus* infection. The polypeptides can also be used in a kit for the immunodetection of *S. aureus* in a sample. *S. aureus* is implicated in numerous human diseases, including cellulitis, eyelid infections, food poisoning, osteomyelitis, skin and surgical wound infections, scalded skin syndrome, toxic shock syndrome, etc. Organisms transformed with the DNA sequences can be used for recombinant production of the polypeptides. The new DNA sequences (and their fragments), are useful as primers or probes for isolating homologues of any of the *S. aureus* DNA sequences contained on the computer readable medium.

Sequence 668 BP; 216 A; 83 C; 119 G; 190 T; 0 U; 60 Other;

Db 242 TTGAATCAGCTTAAAAATATCGGTGATATTGTTGACCGTGATGTTGGCAAGATCT 301  
QY 310 ArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAsp 329  
Db 302 CGTGGCTTAATTTCTTCGAGAAATGCGTAAAGAACTTGTGAATTATCAAAAAAGAA 361  
QY 330 SerLys 331  
Db 362 CAAAAG 367  
RESULT 40  
ID AAS77335  
XX AAS77335 standard; cDNA; 2799 BP.  
AC AAS77335;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #13139.  
XX  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
XX Homo sapiens.  
XX WO200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US008631.  
XX  
XX 31-MAR-2000; 2000US-00540217.  
XX  
XX 23-AUG-2000; 2000US-00649167.  
XX  
XX (HYSE-) HVSEQ INC.  
XX  
XX Dmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
XX  
XX P-PSDB; ABG13148.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
XX Claim 1; SEQ ID NO 13139; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 2799 BP; 619 A; 740 C; 776 G; 664 T; 0 U; 0 Other;  
SQ

Alignment Scores:

Pred. No.: 9,06e-24 Length: 2799  
Score: 350.00 Matches: 69  
Percent Similarity: 68.92% Conservative: 33  
Best Local Similarity: 46.62% Mismatches: 46  
Query Match: 20.71% Indels: 0  
DB: Gaps: 0  
US-10-724-972a-6352 (1-335) x AAS77335 (1-2799)  
QY 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64  
Db 1672 GTTCAGAGCAGACACCGCAGCTTTACACCTGAAAAAACCCACACGATTTGGTGCTG 1731  
QY 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84  
Db 1732 GAACTCTCGTTTCGCGATGCGCTGGCGCGCGTCCGATCCGATTCGGTATTGCCGAC 1791  
QY 85 AspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSer 104  
Db 1792 GATAACCGATGCAAAACGCATCCTGCCCGAAGTGGTCCGCACCTGAAACCGTGGCAGTCC 1851  
QY 105 ValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLysLysProAspLeuIle 124  
Db 1852 GTCGAAACGGCGCGCAGCGAGCCTGGAGCCATTGCCCGCTCTGAAACAGACCTGATC 1911  
QY 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144  
Db 1912 ATTGCCGACAGCAGTCGCCATCGGGGGTTTACATCGCTTGCAGCAATCGCCCGGTA 1971  
QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164  
Db 1972 CTGCTGCTTAAGTCCCGCAACCAACCTACGCTGAAAATTTGCAATCTCGCGGTATCATC 2031  
QY 165 SerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAspLysLys 184  
Db 2032 GCGCAATATGGTGGGTAAAGCGAGAGATGCAGCAGCTCTGGAAACAACATAAAGAGAGG 2091  
QY 185 IleGluGluTyrLysLysGluIle 192  
Db 2092 ATGGCGCAGTGGCGCAGCAGCTT 2115

Search completed: November 9, 2005, 19:55:15  
Job time : 749 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 9, 2005, 19:55:21 ; Search time 3860 Seconds  
(without alignments)  
3303.505 Million cell updates/sec

Title: US-10-724-972A-6352  
Perfect score: 1690  
Sequence: 1 GVESVGLKILSVIGLLFVL.....EEMAKELVLSKKDKKONK 335

Scoring table:  
BLOSUM62 Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 34239544 seqs, 19032134700 residues  
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-O=/cgn2.1/USPTO.spool\_h/US10724972/runat\_03112005\_141900\_29152/app\_query.fasta\_1.519  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10724972 @CNG 1.1 3437 @runat\_03112005\_141900\_29152 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBIOCK=100 -LONGLOG  
-DSV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: gb\_est1:  
2: gb\_est2:  
3: gb\_hic:  
4: gb\_est3:  
5: gb\_est4:  
6: gb\_est5:  
7: gb\_est6:  
8: gb\_gsal:  
9: gb\_gsa2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	412	24.4	776	9	CL694436
c 2	410	24.3	801	9	CL674627
c 3	368	21.8	625	7	CF307287
c 4	210	12.4	531	2	BE432885
c 5	152	9.0	760	9	CL682049
c 6	129	7.6	966	9	CNS01MUR
c 7	126.5	7.5	2019	9	CL958706
c 8	122	7.2	810	9	CL65859
c 9	122	7.2	881	9	CL666398

10	121.5	7.2	4137	9	AY420057
c 11	119.5	7.1	764	4	BJ347654
c 12	119	7.0	839	8	AZ532836
c 13	119	7.0	906	8	AZ538305
c 14	119	7.0	1782	8	CL957361
c 15	118.5	7.0	897	8	BH132864
c 16	118.5	7.0	2740	3	AK029960
c 17	117.5	7.0	1005	5	BQ065769
c 18	117	6.9	1065	7	CF109973
c 19	117	6.9	5504	9	AY407955
c 20	116.5	6.9	804	7	CO573945
c 21	116.5	6.9	832	4	BJ714899
c 22	116.5	6.9	1545	3	AK012454
c 23	116	6.9	816	7	CO113362
c 24	116	6.9	889	8	AZ549761
c 25	116	6.9	1483	3	CR650378
c 26	115.5	6.8	673	1	AL854856
c 27	115.5	6.8	684	5	BW253010
c 28	115.5	6.8	690	5	BW254793
c 29	115.5	6.8	692	5	BW222839
c 30	115.5	6.8	700	5	BW235942
c 31	115.5	6.8	702	5	BW248137
c 32	115.5	6.8	713	5	BW441129
c 33	115.5	6.8	753	5	BW055754
c 34	115.5	6.8	981	4	BI905825
c 35	115.5	6.8	4107	9	AY420059
c 36	115	6.8	785	7	CR426562
c 37	115	6.8	1120	3	CR731171
c 38	115	6.8	2252	3	CR685229
c 39	114.5	6.8	707	4	BJ706325
c 40	114.5	6.8	723	4	BJ714400
c 41	114.5	6.8	727	4	BJ712865
c 42	114.5	6.8	733	4	BJ707678
c 43	114.5	6.8	751	4	BJ712289
c 44	114.5	6.8	811	4	BJ714906
c 45	114.5	6.8			

ALIGNMENTS

RESULT 1  
CL694436/c  
LOCUS  
DEFINITION  
CL694436  
pacificus var. California Pristionchus pacificus genomic, genomic survey sequence.  
CL694436.1 GI:50216344  
GSS.  
Pristionchus pacificus  
Pristionchus pacificus  
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus.  
REFERENCE  
1 (bases 1 to 776)  
Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.  
AppADB: an AcedB database for the nematode satellite organism Pristionchus pacificus  
Nucleic Acids Res. 32 (1), D421-D422 (2004)  
CONTACT: Sommer RJ  
Evolutionary Biology  
Max-Planck-Institute for Developmental Biology  
Spemannstr. 37-39, Tuebingen D-72076, Germany  
Tel: 00497071601371  
Fax: 00497071601498  
Email: ralf.sommer@tuebingen.mpg.de  
This library was generated at Caltech, Pasadena, USA and end sequenced at Vancouver, Canada.  
Seq primer: 17  
Class: fosmid ends.  
Location/Qualifiers  
1..776  
/organism="Pristionchus pacificus"  
/mol\_type="genomic DNA"



```

QY 185 IleGluGluTyrLysLysGluLeuThrMetAspLysAsnGlnLysValLeuProAlaVal 204
   ::::::::::: ::::: ::::: ::::: ::::: :::::
DB 499 ATGGCCACATGGGCCAGCCAG-----CTTCCCAAAGGACACGCGTGGCTTTGGCACA 552

QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224
   ::::: ::::: ::::: ::::: ::::: :::::
DB 553 TCACGGCAACAGCAATTAACCTGCATACTCAGGACACTGGACGGCAGCGTGGCC 612

QY 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu 244
   ||||| ::::: ::::: ::::: ::::: :::::
DB 613 TCTCTGGGGCTG-----AACGTTCCCGCTGGGATG 642

QY 245 LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263
   ||| ::::: ::::: ::::: ::::: :::::
DB 643 GCGGGTGCCTCCATCGCGCTGGAGCAACTGCTGCGGCTCAATCTGCTGG 702

QY 264 MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGlyLys 283
   ::::: ::::: ::::: ::::: :::::
DB 703 CTGCTGGTGGCCACTATCGC-----GAAGAGAGCATGTTAAACGCTGGCAACAA 753

QY 284 AspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgVal 298
   ||||| ::::: ::::: ::::: :::::
DB 754 GATCCGCTCTGGCAGATGTTAACCGCGCGCAGAACACGAGTT 798

```

## RESULT 3

CF307287/c  
LOCUS  
DEFINITION HDAL--06-E11.g1 OSHDAC1-overexpressing transgenic rice lambda phage CDNA library I (HDAL) Oryza sativa (japonica cultivar-group) cDNA clone HDAL--06-E11, mRNA sequence.

## ACCESSION

CF307287.1 GI:33679048

## VERSION

EST.

## KEYWORDS

Oryza sativa (japonica cultivar-group)

## SOURCE

Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzoideae; Oryza.

## REFERENCE

1 (bases 1 to 625)  
Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,

Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.

Large-scale Sequencing Analysis of Rice ESTs

Unpublished (2003)

Contact: Nahm B.H.

Genomics and Genetics Institute, GreenGene Biotech Inc.; Division

of Bioscience and Bioinformatics, Myongji University

Yongin, Kyeonggi, Korea

Tel: 82 31 330 6193

Fax: 82 31 321 6355

Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.

## FEATURES

Location/Qualifiers

1..625 /organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="Nackdong"

/db\_xref="taxon:39947"

/clone="HDAL--06-E11"

/tissue\_type="callus"

/dev\_stage="proliferated callus on 2N6 media for 2 weeks"

/lab\_host="E.coli SOLR"

/clone\_lib="OSHDAC1-overexpressing transgenic rice lambda

phage CDNA library I (HDAL)"

/note="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2:

XhoI; Callus was treated with ABA(20um) for 1hour. cDNA

was inserted into lambda Uni-ZAP XR vector at 5' end with

EcoRI and 3' end with XhoI site. mRNA was derived from

rice Histone Deacetylase overexpression line."

## ORIGIN

Alignment Scores: 9.15e-31 Length: 625  
Pred. No.: 369.00 Matches: 76  
Score:

Percent Similarity: 62.84% Conservative: 39  
Best Local Similarity: 41.53% Mismatches: 66  
Query Match: 21.78% Indels: 2  
DB: 7 Gaps: 1

US-10-724-972A-6352 (1-335) x CF307287 (1-625)

```

QY 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64
   ::::: ::::: ::::: ::::: ::::: :::::
DB 619 GTTCAGGACGACACACGCGCATGTTTACACTCGAAAAAAGCCCAACGATTTGTGTG 560

QY 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84
   ||||| ::::: ::::: ::::: ::::: :::::
DB 559 GAACTCTCTGTTCCCGATGCGCTGCCCGCGGTGGACGTCATCCCATCGTATTGGC 500

QY 85 AspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSer 104
   ||||| ::::: ::::: ::::: ::::: :::::
DB 499 GATAACGATGCAAAACGCGCATCTCGCGAAGTGGCTGGCGACCTGAAACCGGCGAGTCC 440

QY 105 ValGlyThrArgLysGlnProAsnLeuGluGluLysSerLysLeuLysProAspLeu 124
   ||||| ::::: ::::: ::::: ::::: :::::
DB 439 GTCGGAACGCGCGCGCAGCGCGCATTTGCCGCTCTGAAACACGACCTGATC 380

QY 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144
   ||||| ::::: ::::: ::::: ::::: :::::
DB 379 ATTGCCGACGAGCTGCCCATCGCGGGGTTTACATCGCTTGCACAAATCGCGCGGTA 320

QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164
   ::::: ::::: ::::: ::::: :::::
DB 319 CTGTGCTTTAAGTCCCGCAACGAAACCTACGCTGAAAAATTTGCAATCTCGCGCTATCATC 260

QY 165 SerLysAlaLeuGlyLysGluGluGluLysLysArgLeuGluGluHisAspLysLys 184
   ::::: ::::: ::::: ::::: :::::
DB 259 GCGCAATGCTGGGTAAGAACGAGAGATGCAGGCACGCTCTGGAACAACATAAGAGAG 200

QY 195 IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204
   ::::: ::::: ::::: ::::: :::::
DB 199 ATGGCGCAGTGGCGCAGCCAG-----CTTCCAAAGGACACGCGTGGCTTTGGCACA 146

QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224
   ::::: ::::: ::::: ::::: :::::
DB 145 TCACCGCAACAGCAATTAACCTGCATACTCAGGAGACCTGGACCGGCGCTGTGGCC 86

QY 225 GlnLeuGly 227
   |||||
DB 85 TCTCTGGGG 77

```

## RESULT 4

BE432885

LOCUS

DEFINITION

clone cLEG10K14, mRNA sequence.

ACCESSION

BE432885

VERSION

BE432885.1

KEYWORDS

SOURCE

ORGANISM

LYCOPERSICON ESCULENTUM (TOMATO)

LYCOPERSICON ESCULENTUM

EUKARYOTA; VIRIDIPLANTAE;

SPERMATOPHYTES; MAGNOLIOPHYTES;

ASTERIDS; LAMIIDS; SOLANALES;

SOLANACEAE; SOLANUM; LYCOPERSICON

1 (bases 1 to 531)

REFERENCE

AUTHORS

Generation of ESTs from tomato fruit tissue, breaker stage

Unpublished (2000)

Contact: CUGI

Clemson University

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

FEATURES  
source

Location/Qualifiers  
1. .531  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cLEG10K14"  
/tissue\_type="Pericarp"  
/dev\_stage="breaker"  
/lab\_host="SOLR"  
/clone\_lib="tomato breaker fruit, TIGR"  
/note="Vector: pBluescriptSKmCuadapt; Site 1: EcoRI;  
Site 2: XhoI; Fruit were harvested at the breaker stage  
(first sign of lycopene accumulation on the blossom end of  
the fruit). Fruit were cut in half and the seeds and  
locules were discarded prior to freezing the pericarp."

## ORIGIN

Alignment Scores:  
Pred. No.: 4.28e-13 Length: 531  
Score: 210.00 Matches: 41  
Percent Similarity: 75.3% Conservative: 11  
Best Local Similarity: 59.4% Mismatches: 17  
Query Match: 12.43% Indels: 0  
DB: 2 Gaps: 0

US-10-724-972A-6352 (1-335) x BE432885 (1-531)

QY 67 SerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAspAsn 86  
|||||  
Db 4 TCGTTCGGCGATCGCGTGGCCCGCGTGGACGTCATCCCGATCGGTATTGCCGACGATAAC 63  
QY 87 LysLysAsnArgIleLysProLeuArgAspLysIleGlyLysThrSerValGly 106  
|||||  
Db 64 GATGCAAAACGATCTCTGCCGAAGTGGTGGCCACCTGAACCGTGGCAGTCCGTCGA 123  
QY 107 ThrArgLysGlnProAsnLeuGluLysSerLysLeuLysProAspLeuIleAla 126  
|||||  
Db 124 ACGCGCGCGACGCGAGCTGGAAGCATTGCGCTCTGAACACGACCTGATTCG 183  
QY 127 AspAsnAsnArgHisLysGlyIleTyr 135  
|||||  
Db 184 GACAGCAGTCGCGATCGCGGGGTTTAC 210

## RESULT 5

CL682049 760 bp DNA linear GSS 09-JUL-2004  
LOCUS PRI0133d\_D03\_2 - PRI0133d.BR (760) Mixed stage foemid library of P.  
DEFINITION pacificus var. California Pristionchus pacificus genomic, genomic  
survey sequence.

ACCESSION CL682049.1 GI:50189271

VERSION GSS.

KEYWORDS Pristionchus pacificus

SOURCE Pristionchus pacificus

ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;  
Neodiplogasteridae; Pristionchus.

REFERENCE 1 (bases 1 to 760)

Authors Srinivasan, J., Otto, G.W., Kahlow, U., Geisler, R. and Sommer, R.J.

Title AppDB: an AcedB database for the nematode satellite organism

Prionchus pacificus

Nucleic Acids Res. 32 (1), D421-D422 (2004)

Journal Contact: Sommer RJ

Comment Evolutionary Biology

Max-Planck-Institute for Developmental Biology

Spemannstr. 37-39, Tuebingen D-72076, Germany

Tel: 00497071601371

Fax: 00497071601498

Email: ralf.sommer@tuebingen.mpg.de

This library was generated at Caltech, Pasadena, USA and end

sequenced at Vancouver, Canada.

Seq primer: T7

Class: foemid ends.

FEATURES  
source

Location/Qualifiers  
1. .760  
/organism="Pristionchus pacificus"  
/mol\_type="genomic DNA"  
/strain="California"  
/db\_xref="taxon:54126"  
/clone\_lib="Mixed stage foemid library of P. pacificus  
var. California"  
/note="Vector: pEpifos-5 Foemid vector"

## ORIGIN

Alignment Scores:  
Pred. No.: 2.36e-06 Length: 760  
Score: 152.00 Matches: 62  
Percent Similarity: 42.91% Conservative: 50  
Best Local Similarity: 23.75% Mismatches: 111  
Query Match: 8.99% Indels: 38  
DB: 9 Gaps: 11

US-10-724-972A-6352 (1-335) x CL682049 (1-760)

QY 60 ArgValValLeuLeuGlyTyrSerPheValAspAlaLeuValAlaLeuAspValLysPro 79  
|||||  
Db 3 CGTATTGTTCCACGAGCGTCACCGCTGACCGGCTCACTGCTGGCGATTGATGCTCCGGTG 62  
QY 80 ValGlyIleAlaAspAsnLysLysLysArgIleLysPro-----LeuArg 96  
|||||  
Db 63 ATCGCCAGCGCGCGACCCCGCGATTAACCGGTGCGGGATGACCGGGCTTTTACGC 122  
QY 97 -----AspLysIleGlyLysTyrThrSerValGlyThrArgLys----- 109  
|||||  
Db 123 CAGTGGAGCAAGTGGCGAAA-----GAACGCAAACTGCAACGGCTCTAT 167  
QY 110 -----GlnProAsnLeuGluLysSerLysLysLysProAspLeuIleLe 125  
|||||  
Db 168 ATCGCGCAACCGCGCGAGCGTGGCCGCAAAATGCGCGATCTGATTTTAAATTAGC 227  
QY 126 --AlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144  
|||||  
Db 228 GCAACCGCGGGGATTGCGCGCTGGCAGTGTATCATCAGCTTTCATATCGCCCGACA 287  
QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164  
|||||  
Db 288 TTAATCATCAATTACGACGACAAAAAGCTGGCAGTGTGTTAACGCGAGCTTGGCGAAATT 347  
QY 165 SerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAspLysLys 184  
|||||  
Db 348 ACC-----GGCATGAGAAAAAAGCGCGCAGCGGATTCACAGTTTGTATAAGCAA 398  
QY 185 IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204  
|||||  
Db 399 CTGGCAGCGCGGAAGAGCAATCAATTACCGCGCGAG-----CCGCTCACT 446  
QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsn-----SerTyrVal 219  
|||||  
Db 447 GCCATTGCTATACCGCGCTGCACACAGTCCCAATCTCTGCAGCCGGAATCAGCACAA 506  
QY 220 GlyGlnPheLeuSerGlnLeuGlyPheLysGluAla---LeuSerAspAspValThrLys 238  
|||||  
Db 507 GGGCAGATGCTGGGAACAACCTCGGCTTACGCTGGCGAAGTTGCCCGCAGGGGTAAACGCC 566  
QY 239 GlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLysSerGln 258  
|||||  
Db 567 AGCCAAAGCCAGGGCAACCCATGACATCATTCAGCTTGTGGGGAATAATCTGGCTCGG 626  
QY 259 ---ValAsnProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSer 277  
|||||  
Db 627 GGGTTAAACGGTGAATCGCTATTCTCTGTC-----GCCGGTGTATCAGAAAGAC 674  
QY 278 LeuLysGluLeuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 297  
|||||  
Db 675 GCCGATGCGATTATGCAATCCACTGCTCGCGCACCTGCTCGCGCAGTACAAACAAAGCAG 734

```

QY 298 Val 298
Db 735 GTT 737

RESULT 6
LOCUS CNS01MUR 966 bp DNA linear GSS 14-JUN-2001
DEFINITION Anopheles gambiae GSS T7 end of clone 22J04 of NotreDamel library
from strain PEST of Anopheles gambiae (African malaria mosquito),
genomic survey sequence.
ACCESSION AL151412
VERSION AL151412.1 GI:7011891
KEYWORDS GSS.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Anopheles.
REFERENCE 1 (bases 1 to 966)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 966)
AUTHORS Roth,C.W., Brey,P.T., Ke,Z., Collins,F.H. and Weissenbach,J.
TITLE Direct Submission
JOURNAL Submitted (16-FEB-2000) BBMI, Institut Pasteur, 25, rue du Dr.
Roux, Paris 75015, France
COMMENT This clone is from an A. gambiae BAC library provided by F.H.
Collins and sequenced by Genoscope in collaboration with the
Laboratory of Biochem. and Biol. Molec. of Insects, Institut
Pasteur.

FEATURES
source
1..966
Location/Qualifiers
/organism="Anopheles gambiae"
/mol_type="genomic DNA"
/strain="PEST"
/db_xref="taxon:7165"
/clone="22J04"
/clone_lib="NotreDamel"
/note="end : T7"

ORIGIN
Alignment Scores:
Pred. No.: 0.00128 Length: 966
Score: 129.00 Matches: 67
Percent Similarity: 43.21% Conservative: 54
Best Local Similarity: 23.93% Mismatches: 116
Query Match: 7.63% Indels: 44
DB: 9 Gaps: 11

US-10-724-972a-6352 (1-335) x CNS01MUR (1-966)

QY 44 GluileYHieGluGlyThrLysValProLYsHieProLYsArgValValVal 63
Db 160 CAGGTGACCGACAGCGCGGGTGATACCTTGAGCAACAACCGACGGCATGCTCC 219
QY 64 LeuGluTYrSerPheValhapAlaLeuAlaLeuAspVallyProValGly-IleAl 83
Db 220 ACCAGCGTGCATTAACCGGCTCCCTGCTGGCCATTRACGCACCGGTCGCCACGGGS 279
QY 83 aAspAspAenLYsAsnArg-----IlelleLYsProLeuArgAspLY 98
Db 280 CGACACCGCGGAACAACCGGTGGCGNATGGACAGGGTTCTCGGCCCA----- 328
QY 98 sIleGlyLYsTYrThrSerValGlyThrArgLYs-----GlnPr 111
Db 329 -GTGGGGCGCATATGGCGAA-----CARCGCAAGTTGGCCGCTGTATATCGCGAGCC 381
QY 111 oAnLeuGluGluIleSerLYsLeuLYsProAspLeuIlelleAlaAspAsnArgHi 131
Db 382 GAGCGTGAAGCGGTGGCGCGCAAAATTCGGACCTKATTTTTTAT-CAGCGCACCGGSG 440

```

```

QY 131 sLYs-----GlylleTYrLYsAspLeuAsnLYsIleAlaProThrIleGluLeuLY 148
Db 441 GGATTCCCGCGCTGGCGCTGTACGATCAGCTTTCCGCCATTGCGCGAGCTTATCATCAA 500
QY 148 sSerPheAspGlyAspTYrAsnGluAsnIleAspAlaPheLYsThrIleSerLYsAlaLe 168
Db 501 CTACKACCAAGAGAGCTGGCAGGTGCTCTGACCAGCTGGCGGAGATCACC----- 552
QY 168 uGlyLYsGluGluGlyLYsLYsArgLeuGluGluHisAspLYsLYsIleGluGluTY 188
Db 553 -GGACACGAAAGCAGCGCGCAGCGCATTTGATTAACAGCTCGCGCAGGT 611
QY 188 rLYsLYsGluIleThrMetAspLYs-----AsnGlnLYsValLeuProAlaValAl 205
Db 612 GAAACAGCAGATGACGCTGCCCGCCGACCGCGTGAACGCCATCGTCTACACCGCGCGCC 671
QY 205 aAlaLYsSerGlyLeuLeuAlaHisProSerAsnSerTYrValGlyGlnPheLeuSerGl 225
Db 672 GCATTCCCGCCAACTTTTGACC-----ACGGAATCGGCACAGGCAAGCTCTGCACCA 725
QY 225 nLeuGlyPheLYsGluAlaLeuSerAspAspValThrLYsGlyLeu-----Se 241
Db 726 GCTGGCTTTAGCGTGCG-----GACCTGCTGCGCGGTTGCAGACCTCAACAAG 776
QY 241 rLYsTYrLeuLYsGlyProTYrLeuGlnMetAsnThrGluThrLeuSerGln--ValAsn 260
Db 777 CCAGGGCAAGCGCCACGACATTTATCCAGCTGGCGGTGAAACCTKGCACGGGGTGAAC 836
QY 261 ProGluArgMetPheIleMetThrAsnLYsAlaSerSerAsnGluProSerLeuLYsGlu 280
Db 837 GCGGAAGCGCTGTTTKTGT-----GCCGTGACGAGAAAGACCTKCGCGCG 884
QY 281 LeuGluLYsAspProValTrpLYsLYsLeuAsnAlaValLYsAsnGlnArgVal 298
Db 885 ATTTAGCTAACCCGCTGCTKGCACCTGCGCTGCGTAAATAACAAGCGCGTA 938

RESULT 7
LOCUS CL958706 2019 bp DNA linear GSS 21-SEP-2004
DEFINITION OaIFCC001947 Oryza sativa Express Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.
ACCESSION CL958706
VERSION CL958706.1 GI:52372164
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 2019)
AUTHORS Ma,L., Wang,C., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
Wong,G.K.S., Deng,X.W. and Wang,J.
TITLE An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
JOURNAL Unpublished (2004)
COMMENT Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
FEATURES
source
1..2019
Location/Qualifiers
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences"

```

## ORIGIN

Alignment Scores:  
Pred. No.: 0.00686 Length: 2019  
Score: 126.50 Matches: 74  
Percent Similarity: 39.33% Conservative: 55  
Best Local Similarity: 22.56% Mismatches: 114  
Query Match: 7.49% Indels: 85  
DB: 9 Gaps: 13

US-10-724-972A-6352 (1-335) x CL658706 (1-2019)

QY 37 GluSerSerLysAspGlyValGluLeuLysHisGluGluGlyThrThrLysValProLys 56  
Db 1162 CAACCAACAAAGGATTAGTGAGGTAGAGATGATAAACTTCGGTGGATGTTTATCT 1221

QY 57 HisProLysArgValValLysLeuGluTyrSerPhe-----ValAsp 70  
Db 1222 GCACAAACAAAGCATATTAGAAGTAGAGGCTAACAAATCTCAAGATGGAATGGGT 1281

QY 71 AlaLeuValAlaLeuAspValLysProValGlyIle---AlaAspAsnLysLysAsn 89  
Db 1282 GCATTTGGTG-----GAGCAAGAGGTTGGCTACGAAGGCATTTCGCGCTGAGAAGCA 1335

QY 90 ArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLys 109  
Db 1336 AAAATAACGAAGGATTCGAGGATGTTAAGAGAGAGATGGAGAAATCCAGGTCAAGAG 1395

QY 110 GlnProAsnLeuGluGluLysSerLysLysLysProAspLeuIleAlaAspAsnLys 129  
Db 1396 GATTGGTGAGGTCGAAAGGATTAAGCTTCGGTTGGAGATATTAAATAGCGGAGCAAAA 1455

QY 130 ArgHis-----LysGlyIleTyrLysAspLeuAsnLysIleAlaPro 143  
Db 1456 CATACGATGCTGAATGGAGGTTAAAGGCTCAAGATGGAACCTTGGTGCATTGGCAGAG 1515

QY 144 ThrIleGlu-----LeuLysSerPheAsp----- 151  
Db 1516 GCAATAGACAGCTGTGAAGTCATTGATGCTGAGAGGAGAAATTCATCAGGGAATG 1575

QY 152 GlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGlu 171  
Db 1576 GGGGACCTCAAGAGGAAGATAGAGAAATCCAGGTCAGCAAGGAAGCT-----GCTGAG 1629

QY 172 GluGluGlyLysLysArgLeuGluHisAspLys---LysIleGluGluTyrLysLys 190  
Db 1630 GAAGTAGGCGCAACAGGAATCCAGAGGCTGATAGCTGAGGCTGAGTGGTGAAGATC 1689

QY 191 GluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaAlaLysSerGlyLeu 210  
Db 1690 CAAGTTCTTTATCGCAACTACAGCATCTTATATGAACCTTGATGCTAAGCATTCCTCTC 1749

QY 211 LeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGlu 230  
Db 1750 CTAATATGATGAGAAGATTCA-----GCCAGAGAA 1779

QY 231 AlaLeuSerAspAspValThrLysGlyLeuSerLysTyrLysLysGlyProTyrLeuGln 250  
Db 1780 GCATTTA-----GATGTTGAGAAGGTTGAAGCATTAAGCTGAAGTCAAAATTTAAGGAA 1833

QY 251 MetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsnLys 270  
Db 1834 CTTGAAAT----- 1842

QY 271 AlaSerSerAsnGluProSerLeuLysGluLeuGluLysAspProValTyrLysLysLeu 290  
Db 1843 -----TATAAGGCTGAAAAGGATGAAGAGGCTGGGAAACTG 1878

QY 291 AsnAlaVal-----LysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAla 307  
Db 1879 AAGCGGCGCATTCGAGAAAGAAAGTAGTATGATGTCCTCATCAAGGAC----- 1929

QY 308 ArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLys 327

Db 1930 -----AATGAGTTGCTCGGTCTCGCAATT 1953  
|||||::: |||::: |||:::

QY 328 LysAspSerLysLysAspAsnLys 335  
Db 1954 OCTGAGGCACAGGAAAGAAATAAG 1977  
:::|||||

## RESULT 8

CL65859

## LOCUS

DEFINITION

CL65859

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Max-Planck-Institute for Developmental Biology

Spemannstr. 37-39, Tuebingen D-72076, Germany

Tel: 00497071601371

Fax: 00497071601498

Email: ralf.sommer@tuebingen.mpg.de

This library was generated at Caltech, Pasadena, USA and end

sequenced at Vancouver, Canada.

seq primer: 17

Class: fosmid ends.

Location/Qualifiers

1. 810

/organism="Pristionchus pacificus"

/mol\_type="genomic DNA"

/strain="California"

/db\_xref="taxon:54126"

/clone\_lib="Mixed stage fosmid library of P. pacificus

var. California"

/note="vector: pEpifos-5 Fosmid vector"

ORIGIN

Alignment Scores:

Pred. No.: 0.00611 Length: 810

Score: 122.00 Matches: 29

Percent Similarity: 48.25% Conservative: 26

Best Local Similarity: 25.44% Mismatches: 45

Query Match: 7.22% Indels: 14

DB: 9 Gaps: 3

US-10-724-972A-6352 (1-335) x CL65859 (1-810)

QY 213 HisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232

Db 35 CATACTCAGGAGACCTGGACCGCAGCGCTGCTGCGCTCTCTCTGGGCTG----- 82

QY 233 SerAspAspValThrLysGlyLeuSerLysTyrLysLysGlyProTyrLeu---GlnMet 251

Db 83 -----AACGTTCCCGCTCGGATGGCGGTGGTCCATCCCGTCCATC 124

QY 252 AsnThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsnLysAla 271

Db 125 GGCCTGGAGCAACTGCTGGCGGTCAATCTCGCTGGCTGCTGGTGGTCCACTATCGC--- 181

QY 272 SerSerAsnGluProSerLeuLysGluLeuGluLysAspProValTyrLysLysLeuAsn 291

Db 182 -----GAAGAGAGCATTTGTTAAACGCTGGCAACAGATCCGCTCTGGCAGATGTTAAC 235





Db 3076 AAAACAGACGAGCTAAACAAACAGTTAAAGAGCTTGTCCACAGAAATACACGGAAGTAAAG 3135  
Qy 107 ThrArgLysGlnProAsnLeuGluGluLeuSerLysLeuLysProAspLeuLeuLeuLeu 126  
Db 3136 AATGTGAAGAGAGAGCTAGTAGAAGAAATGCAACACAGACTTCTGAGATACTTCCAGATG 3195  
Qy 127 AspAsn-----AsnArgHis-----LysGlyLeuLysLysAsp 137  
Db 3196 CAAATCTTTTGCACAAACACATGTTCCATTGGACAGAGTTGAGGCTCTGAAAAATCT 3255  
Qy 138 LeuAsnLysIleAlaProThrIle-----GluLeuLysSerPheAspGlyAspTyrAsn 155  
Db 3256 CTTAATGGCACAATTTCAAAATCTAAAGGAAGAACTGAAGAGTATGCAAGAGTGTACGAG 3315  
Qy 156 GluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGluLys 175  
Db 3316 AAAGAGCAGCAGACAGTACCAACATGTCATCAATTTGTTGGAGAAATCAAAAGAACTCTTCT 3375  
Qy 176 LysArgLeuGluGluHisAspLysLysIleGluGluTyrLysLysGluLeuLeuMet--- 194  
Db 3376 GTACCCCTGGCAGACATTTGAGATTAAAGAGCATTTGAGAAAGAGTTGGATCATATA 3435  
Qy 195 -----AspLysAsnGlnLysValLeuProAlaValAlaLysSerGly 209  
Db 3436 AAAGCCAGCTTGAGAGAAAGGAAGAAAGAACCCAAACAAATGGAAGAACTCTCCAA 3495  
Qy 210 LeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLys 229  
Db 3496 CTTCACTCGGAGGTTTCAGATACT-----AAACAAAGCATTTAAAAAATTTAGAGACTAGA 3549  
Qy 230 GluAlaLeuSerAspValThrLysGlyLysLysValLysLysLysGlyProTyrLeu 249  
Db 3550 GAGGTAGTTGAC-----TTGTCTTAATAT---AAAGCAACAAAGT 3588  
Qy 250 GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn 269  
Db 3589 GATTTGGAGACAG---ATTCTAGCTTAAT---GAAAAATTTGCCAATCTGAATAGA 3642  
Qy 270 Lys-----AlaSerSerAsnGluProSerLeuLys 279  
Db 3643 AAGTATGAGGAAGTATGTGAGGAAGTTTTCATGTCACAAAGAAAGAAATATCTGCACAA 3702  
Qy 280 GluLeuGluLysAspProValTrpLysLysLeu---AsnAlaValLysAsnGln----- 296  
Db 3703 GAT---GAGAGGAATTACTGCTATTCAGCATTTGACAGAGAATTAAGATCAGAGGNA 3759  
Qy 297 ArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArgGlyLeuLysSerSerGlu 316  
Db 3760 CGATGTGAT-----AAGTCCTTAACACAAATCACA 3789  
Qy 317 GluMetAlaLysGluLeuValGluLeuSerLysAspSerLysLysAspAsnLys 335  
Db 3790 GAGTTTACAAAGAAGAATAACAAGATCTCTGCTAAACAAATAGAGCAAGAAAGATAAAG 3846

RESULT 11  
LOCUS BJ347654/c 764 bp mRNA linear EST 06-MAR-2002  
DEFINITION BJ347654 Dictyostelium discoideum cDNA library, AF Dictyostelium  
ACCESSION BJ347654  
VERSION BJ347654.1 GI:19218161  
KEYWORDS EST.  
SOURCE Dictyostelium discoideum  
ORGANISM Dictyostelium discoideum  
REFERENCE Eukaryota; Mycetozoa; Dictyostelid; Dictyostelium.  
AUTHORS 1 (bases 1 to 764)  
TITLES Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.  
Full length cDNA of Dictyostelium discoideum at the aggregation  
stage  
JOURNAL Unpublished (2002)  
CONTACT Tadaeu Shin-i  
COMMENT Center For Genetic Resource Information  
National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.  
Location/Qualifiers  
1. 764  
/organism="Dictyostelium discoideum"  
/mol\_type="mRNA"  
/strain="AX4"  
/db\_xref="taxon:44689"  
/clone="dda27b21"  
/sex="mat A"  
/dev stage="Aggregation stage"  
/clone\_lib="Dictyostelium discoideum cDNA library, AF"

## ORIGIN

Alignment Scores:  
Pred. No.: 0.0108 Length: 764  
Score: 119.50 Matches: 48  
Percent Similarity: 45.37% Conservative: 45  
Best Local Similarity: 23.41% Mismatches: 77  
Query Match: 7.07% Indels: 35  
DB: 4 Gaps: 6  
US-10-724-972A-6352 (1-335) x BJ347654 (1-764)  
Qy 22 AlaThrAlaAlaCysGlyAsnAsnSerSerAsnSerSerLysGluSerSerLysAsp 41  
Db 748 TCATCATCATCATCATCATCATCATCATCATCATCATCATCTTTGGAAACAAACAAATGTA 689  
Qy 42 GlyValGluIleLysHisGluGluGlyThrThrLysValProLysHisProLysArgVal 61  
Db 688 GAATTTGGAAATCAACACCAACAAATTTGGAAAGGAATTTGGAAAGGAATTTAC 629  
Qy 62 ValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGly 81  
Db 628 CAAGAAATTTGGAAAGTAAATATACCAAGAA-----TTGGAAACTGTATAACAAAGAAATTTGGA 575  
Qy 82 -----Ile 82  
Db 574 TGTAAAAACCTAGAAATTTAAAAATCAAAACCAAGAAATTTGGAAATTTAATAACAAAGAAATTTG 515  
Qy 83 AlaAspAspAsnLysLysAsnArgIle---IleLysProLeuArgAspLysIleGlyLys 101  
Db 514 GAAATCAAAACCAAGAAATTTGAAATCAAAATCCAGAAATTTAGAAACCATTAATCAATCAA 455  
Qy 102 TyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLysPro 121  
Db 454 TACAAAATGAAATAGTGCCAAAGATATGGAATATTTGATATTAATAATAAAAGAAAGAA 395  
Qy 122 AspLeuIleIleAlaAspAsnAsnArgHisLysGlyIle-----TyrLysAspLeu 138  
Db 394 GAACCTTGAGAGCCATAGAGAATAGAGATGATATTTAAAAAAGGATATTTGGAACTT 335  
Qy 139 AsnLysIleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIle 158  
Db 334 GAAAGGTCAAT-----AAAGAAATCAAAAGAAATTTATGAAAAATCAAT 290  
Qy 159 AspAlaPhe-----LysThrIleSerLysAlaLeuGlyLysGluGluGlu 173  
Db 289 TTTGATTCATGAAGCAAGAAATTTGAAGAAATCAATAAATTAGTAATAATAATAATGAT 230  
Qy 174 GlyLysLysArgLeuGluGluHisAspLysLysIleGluGluTyrLysLysGluLeuThr 193  
Db 229 TCACAAACCAATTTGGGAGAGAAATAATGAATGTGTACCTTTTAAAAAAGAAAGATTGAA 170  
Qy 194 MetAspLysAsnGln 198  
Db 169 ACATTAATAAATGAA 155

RESULT 12  
LOCUS AZ532836  
839 bp DNA linear GSS 03-NOV-2000

```

DEFINITION  ENTDA85TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
              genomic, genomic survey sequence.
ACCESSION   AZ532836
VERSION     AZ532836.1 GI:11088430
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica
            Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
            1 (bases 1 to 839)
REFERENCE   Loftus,B., Van Aken,S. and Fraser,C.
            Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
            Unpublished (2000)
JOURNAL     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@igr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Forward
            Class: shotgun
            High quality sequence start: 21
            High quality sequence stop: 820.
            Location/Qualifiers
              1..839
                /organism="Entamoeba histolytica"
                /mol_type="genomic DNA"
                /strain="HM1:IMSS"
                /db_xref="taxon:5759"
                /clone_lib="Entamoeba histolytica Sheared DNA"
                /notes="Vector: pHO31; Site 1: Bst I; Constructed at The
                Institute for Genomic Research (TIGR), Rockville, MD.
                Genomic DNA isolated from broth cultures of E. histolytica
                using a method described by Clark and Diamond (Clark,
                C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
                method for isolate identification. Exp. Parasitol.
                77:450.). The DNA was mechanically sheared to give a
                tight size distribution (~2 kb). The v + i method used for
                the library construction is described in detail in Smith,
                H.O. and Venter, J.C. (Making small insert libraries for
                whole genome shotgun sequencing projects. In Genome
                Sequencing: A Practical Approach, eds. M. Vaudin and B.
                Borell, Oxford University Press, 1999)."
```

ORIGIN

Alignment Scores:			
Pred. No.:	0.014	Length:	839
Score:	119.00	Matches:	62
Percent Similarity:	41.47%	Conservative:	45
Best Local Similarity:	24.03%	Mismatches:	103
Query Match:	7.04%	Indels:	48
DB:	8	Gaps:	13

US-10-724-972a-6352 (1-335) x AZ532836 (1-839)

```

QY      35 SerLysGluSerSerLysAspGlyValGluIleLysHis-----GluGluGlyThrThr 52
DB      61 TCTAAAGACAGAAAGCAATGATTGATTCATTTAGACAAATGCTTACAGACACATGAA 120
QY      53 LysValProLysHisProLysArgValVal-----
DB      121 AAGTCTTACCAATTTAAAAAGGATTTATGATCTTTAGACAAATGCTGGAACCTGAA 180
QY      63 ----ValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAsp----ValLysProVal 80
DB      181 ACTGCTTTAAACCAAGTACAGTAAAGATTTATTACTGCTCCCTGATGGAATTCGTGAACT 240
QY      81 GlyIleAlaAspAsnLysAsnArgIleIleLysProLeuArgAspLysIle---- 99
DB      241 GGAATGGGAGATTTCAATAAGGAATTCGTGTTGTAGAGAAATCTTTGTGATTGTAT 300

```

```

QY      100 ---GlyLysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLys 118
DB      301 AATGAAGACACACAAAGTATT-----AATCAAGGACAACTAATGCTTTTACTACTCT 351
QY      119 LeuLysProAspLeuIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeu 138
DB      352 CTTAAAAAAGAA-----TACACTGTAGAAAGTAAAGAAAGTA 387
QY      139 AsnLysIle-----AlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGlu 156
DB      388 CAACAGTTATCAACAAGTACACAAAGAAAGAAAGAAATTTGATCAAGAC----- 438
QY      157 AsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGlyLysLys 176
DB      439 -----TTATCTAATGCAGAAAGAAACGGCATCTACTTCTAATTTTAAAGAA 486
QY      177 ArgLeuGluGluHisAspLysLysIleGluGluTyrLysLysGluIleThrMetAspLys 196
DB      487 GCTATGGGAAGATAT---AAACGTATAGAAAGAAAGAAAGAAATTTTAATGAGTCAT 543
QY      197 AsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsn 216
DB      544 TTACAAGTAACGTGTTATGCTTAACAGAGCTACTTTTGGT----- 582
QY      217 SerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspVal 236
DB      583 TCTATTATTGGGTCATTT-----GCTTATTCTTTAAATCTACTGTTAGTTTCATTGACG 636
QY      237 ThrLysGlyLeuSerLysTyrLysLys---GlyProTyrLeuGlnMetAsnThrGluThr 255
DB      637 TCATCTGGA---GATAAGTTCAATTCATATGAGCCCTTGGCTTAATAAATATCAAAAGTCT 693
QY      256 LeuSerGlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSerSer 273
DB      694 AGTCTTATATCCGAGAGAACTACGTGCTTTGATGACAAATATAAAACAAAGT 747

RESULT 13
LOCUS    AZ538305
DEFINITION  ENTPO68TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, genomic survey sequence.
ACCESSION  AZ538305
VERSION    AZ538305.1 GI:11143120
KEYWORDS   GSS.
SOURCE     Entamoeba histolytica
            Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
            1 (bases 1 to 906)
REFERENCE  Loftus,B., Van Aken,S. and Fraser,C.
            Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
            Unpublished (2000)
JOURNAL    Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@igr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Forward
            Class: shotgun
            High quality sequence start: 142
            High quality sequence stop: 861.
            Location/Qualifiers
              1..906
                /organism="Entamoeba histolytica"
                /mol_type="genomic DNA"
                /strain="HM1:IMSS"
                /db_xref="taxon:5759"
                /clone_lib="Entamoeba histolytica Sheared DNA"
FEATURES
  source

```







Qy 204 ValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeu 223

## Alignment Scores:

Pred. No.: 0.0265 Length: 1005  
Score: 117.50 Matches: 61  
Percent Similarity: 37.9% Conservative: 59  
Best Local Similarity: 19.30% Mismatches: 125  
Query Match: 6.95% Indels: 71  
DB: 5 Gaps: 11

US-10-724-972A-6352 (1-335) x BQ065769 (1-1005)

QY 27 GlyAenAenSerSerSerAenSerSerLysGluSerSerLysAspGlyValGluLeuLys 46  
DB 15 GGAGGAAGAACCAAGCCCTGTCCTCGGCGCTTGAAGAGCCCTTGAAGCCCAA 74  
QY 47 HisGlu---GluGlyThrThrLysValProLysHisProLysArgValValLeuGlu 65  
DB 75 GAGGAACCTCGAGCGGACCAACAAA-----ATGCTCAA 107  
QY 66 TyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyLeuAlaAsp 85  
DB 108 GCCGAATGGAAGACTGGTCCAGCTCCAAGGATGCTGGGCAAGACGCTCCATGAGCTG 167  
QY 86 AenLysLysAsnArgIleLeuLysProLeuArgAspLysIleGlyLysTyrThrSerVal 105  
DB 168 GAGAAGTCCAAGCGGCCCTGGAG----- 191  
QY 106 GlyThrArgLysGlnProAenLeuGluGluLeuSerLysLeuLysProAspLeuLeu 125  
DB 192 ---ACCCAGATGGAGAGATGAAGCGCAGCTGGAGAGCTGGAGGAGCTGCAAGCC 248  
QY 126 AlaAspAenAsnArgHisLysGlyLeuLysAspLeuAsnLysIleAlaProThrIle 145  
DB 249 ACGGAGGAC-----GCCAACTCGCGGTGAAGTC 278  
QY 146 GluLeuLysSerPheAspGlyAspTyrAenGluAenIleAspAlaPheLysThrIleSer 165  
DB 279 AACATGCGAGCGCTCAAGCGCAGTTCGAAGGATCTCCAAGCC----- 323  
QY 166 LysAlaLeuGlyLysGluGluGluGlyLysArgLeuGluGluHisAspLysLysIle 185  
DB 324 ---CGGAGCGAGCGAATGAGGAGAGAGAGGCAACTCGCAGAGCAGCTT 371  
QY 186 GluGluTyrLysLysGluLeuThrMetAspLysAenGlnLysValLeuProAlaValAla 205  
DB 372 CACGAGTATGAGACGGAAGCTGGAAGCGGCGAAGCAACGTCCTCGCAGCTGCAGCA 431  
QY 206 AlaLysSerGlyLeuLeuAlaHisProSerAenSerTyrValGlyGlnPheLeuSerGln 225  
DB 432 AAGAAG-----AAGCTGAAGGGGAC 452  
QY 226 LeuGlyPheLysGluAlaLeuSerAspAspValThrLysGlyLeuSer-----Lys 242  
DB 453 CTGAAGACCTGGAGCTTCAGCGGACTCTGCATCAAGGGGAGGAGGAAGCCATCAAG 512  
QY 243 TyrLeuLysGlyProTyrLeuGlnMetAen-----ThrGluThrLeuSerGlnValAen 260  
DB 513 CAGCTACGCAAACTCGAGCTCAGATGAAGACTTTCGAAGAGACTGGAAGATCCCGT 572  
QY 261 ProGluArgMetPheIleMetThrAenLysAlaSerAenGlnProSerLeuLysGlu 280  
DB 573 GCCTCCAGAGATGAGATCTTTGCCACA---GCCAAGAGATGAGAAGAAAGCCAAAGAGC 629  
QY 281 LeuGluLysAspProValTrpLysLysLeuAenAlaValLysAenGlnArgValAspIle 300  
DB 630 TTGAAGCAGACCTCATGCAGCTCAAGAGGACCTCGCGCGCTGAGAGG----- 680  
QY 301 LeuAspArgAspLeuTrpAlaArgSerArgGlyLeuLysSerSerGluGlnMetAlaLys 320  
DB 681 -----GCTCGCAACAAAGCGGACCTCGAAGAGGAACTGCGCAGAG 722  
QY 321 GluLeuVal---GluLeuSerLysLysAspSerLysLysAspAsnLys 335  
DB 723 GAGCTGGCCAGTAGCGTGTGCGGAAAGGACGCACTCCAGGACGAGAAG 770

RESULT 18  
CF109973

LOCUS 1065 bp mRNA linear EST 23-JUL-2003  
DEFINITION Shultzomica03224 Rat lung airway and parenchyma cDNA libraries  
Rattus norvegicus cDNA clone Contig2841 5', mRNA sequence.

ACCESSION CF109973

VERSION CF109973.1 GI:33165360

KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)

## ORGANISM

Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 1065)

## AUTHORS

Shultz,M.A., Zhang,L., Gu,Y.-Z., Baker,G.L., Fannuchi,M.V.,  
Padua,A.M., Gurske,W.A., Morin,D., Penn,S.G., Jovanovich,S.B.,  
Plopper,C.G. and Buckpitt,A.R.

TITLE Gene expression analysis in response to lung toxicants: I.

JOURNAL Sequencing and microarray development

COMMENT Am J Respir Cell Mol Biol. 30 (3), 296-310 (2004)

Contact: Shultz MA

Dept. of Molecular Biosciences, School of Veterinary Medicine

University of California, Davis

1311 Haring Hall, One Shields Avenue, Davis, CA 95616, USA

Tel: 530 752 0793

Fax: 530 752 4698

Email: mashultz@ucdavis.edu

Average Phred score is 20 or better. All poor quality data (Phred <  
20) and vector/linker sequence has been removed.

High quality sequence stop: 1065.

## FEATURES

Location/Qualifiers

1..1065

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/strain="Sprague-Dawley"

/db\_xref="taxon:10116"

/clone="Contig2841"

/sex="male"

/tissue\_type="airway or parenchyma"

/dev\_stages="adult"

/clone\_lib="Rat lung airway and parenchyma cDNA libraries"

/note="Organ: lung; Vector: pGEM-11zf(-); Site.1: Eco RI;

Site 2: Not I; mRNA was isolated from microdissected rat

lung airways and parenchyma tissues."

## ORIGIN

## Alignment Scores:

Pred. No.: 0.0327 Length: 1065

Score: 117.00 Matches: 51

Percent Similarity: 37.56% Conservative: 29

Best Local Similarity: 23.94% Mismatches: 59

Query Match: 6.92% Indels: 74

DB: 7 Gaps: 9

US-10-724-972A-6352 (1-335) x CF109973 (1-1065)

QY 28 AenAenSerSerSerAenSerSerLysGluSerSerLys----- 40

DB 413 AACAAATGAAGAAGAAATCTGCCTCAAAACAAGATAAAGAGGCAAAAGGCAAG 472

QY 41 -----AspGlyValGluLeuLysHisGluGlyThrThrLys 53

DB 473 AAAACAAGTTTGTAGTAGAATGTAGTGAAGAAATGGAGATAAAGATCAAAATCAAA 532

QY 54 ValProLysHisProLysArgValValLeuGluTyrSerPheValAspAlaLeuVal 73

DB 533 AAGCTGCCAGACCTTAACCTCAGAGGTGCTCTCT----- 568

QY 74 AlaLeuAspValLysProValGlyLeuAlaAspAenLysLysAsnArgIleLys 93

DB 569 -----GGGAGTGAGGATGCAGATGATCTCTTAATAAACTT----- 601



QY 94 ProLeuArgAspLysIleGlyThrSerValGlyThrArgLysGlnProAsnLeu 113  
Db 602 -----TCTAAAGAGTAAAGAGCCAGCAATCACTAAAGAGCGGATGGTCG 652  
QY 114 GluGluIleSerLysLeuLysProAspLeuIleAlaAspAsnAsnArg----- 130  
Db 653 GAAGAGGATGAA-----GACATAGCAAGAAAGTAA 685  
QY 130 ----- 130  
Db 686 GAACGTTTCGAGAGTAACTCTCAGGTGAAGTGGTGCAGTCTGATCAATTTTTCAG 745  
QY 131 ---HisLysGlyIleThrLysAsp---LeuAsnLysIleAlaProThrIleGluLeuLys 148  
Db 746 TCCAGAAAGGAGCAGAAAATAATCAGAGCAACAGTCAGTCTCCATCAGTACAGTGG 805  
QY 149 SerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSer-----Lys 166  
Db 806 AATGAGGACGATGACTCTTCTTCAAAAT-----AAGACGGTGGCCAGAGAAG 856  
QY 167 AlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAspLysLysIleGlu 186  
Db 857 GCAGAAAGAGAGCGTGCAGAGGAGAAAGAGAG---GAAGAGGAGAGCAAAAGTTGCGA 913  
QY 187 GluTyrLysLysGluIleThrMetAspLysAsnGlnLys 199  
Db 914 AAGGTGAAGAGAGAGAGAGCTAGAGAGGCGAGGAG 952  
AY407955 5504 bp DNA linear GSS 15-DEC-2003  
Homo sapiens MYH11 gene, VIRTUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.  
ACCESSION AY407955  
VERSION AY407955.1 GI:39763926  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,  
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,  
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,  
Adams,M.D. and Cargill,M.  
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,  
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,  
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,  
Adams,M.D. and Cargill,M.  
AUTHORS Direct Submission  
TITLE Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
JOURNAL Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES Location/Qualifiers  
source  
1..5504  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
gene  
c1..>5504  
/gene="MYH11"  
/locus\_tag="HCM3063"  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.328 Length: 5504  
Score: 117.00 Matches: 73  
Percent Similarity: 35.32% Conservative: 63

Best Local Similarity: 18.96% Mismatches: 146  
Query Match: 6.92% Indels: 103  
DB: 9 Gaps: 11  
US-10-724-972A-6352 (1-335) x AY407955 (1-5504)  
QY 22 AlaThrAlaAlaCysGlyAsnAsnSerSerSerLysGluSerSerLysAsp 41  
Db 4137 GCTGCAGGACTTTGCGCAGCACCGTGGAAAGCTCTGGAAGAGGAGAGAGAGGTTCCAGAA 4196  
QY 42 Gly-ValGlu-----IleLysHisGluGlyThrThrLysValProLysHisPr 58  
Db 4197 GGAGATCAGAACTCACCACAGCAGTACAGAGAGAGGCGCGCTTATGATAAAGTGA 4256  
QY 58 oLysArgValValValLeuGluTyrSerPheValAspAlaLeuAlaLeuAspVally 78  
Db 4257 AAGACCAAGAACAGCGCTTCAGCAGGAGCTGGACGACCTGGTGTGTTGATTGCAACCA 4316  
QY 78 sProValGlyIleAlaAspAsnLysLysAsnArgIleLysProLeu----- 95  
Db 4317 GCGCAACTCGTGTCCAACTCGAAAGAGCAGAGGAAATTTGATCAGTTGTTAGCCGA 4376  
QY 96 -----Ar 96  
Db 4377 GGAGAAACATCTCTTCAATACCGGATGAGAGGACAGAGCTGAGGAGAGAGCCAG 4436  
QY 96 sAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluGlu 116  
Db 4437 GGAGAGGAAACCAAGGCCCTGTCTCGCTCGGSCCTTGAAGAGGCGCTTGAAGCCAA 4496  
QY 116 sSerLysLys-----ProAspLeuIleAla 126  
Db 4497 AGAGAACTCGAGCGGACCAAAATGCTCAAAGCCGAAATGGAAGACCTGTCAGCTC 4556  
QY 126 aAspAsnAsnArgHisLysGlyIleTyrLys----- 136  
Db 4557 CAAGATGACGTGGCAAGACGTCTCAGTGTGAGAGAGTCCAAGCGGCGCTGAGAG 4616  
QY 137 -----As 137  
Db 4617 CCAGATGAGGAGATGAAGACGACGCTGGAAGAGCTGGAGGAGCAGCTGCAAGCCAG 4676  
QY 137 pleuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGlu 157  
Db 4677 GGACGCCAACTGCGCTGGAAGTCAACATGAGCGCTCAAGGCGCAGTTTCAAGAGGGA 4736  
QY 157 nIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGlyLysLys 177  
Db 4737 TCTCAAGCC-----CGGACGAGCAGAGATGAGAGAA 4769  
QY 177 gLeuGluGluHisAspLysLysIleGluGluTyrLysLysGluIleThrMetAspLys 197  
Db 4770 GAGGAGCACTGCAGACAGCAGCTTCACAGATATGAGACGGAAGTGGAGAGCAGCGAA 4829  
QY 197 nGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSe 217  
Db 4830 GCAAGCTGCTCGCAGCTGCAGCAAGAG----- 4860  
QY 217 rTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValTh 237  
Db 4861 -----AAGTGAAGGGGACCTGAAAGACCTGAGGCTTACGGCGGCTCTGCCAT 4910  
QY 237 rLysGlyLeuSer-----LysTyrLysGlyProTyrLeuGlnMetAsn----- 252  
Db 4911 CAAGGAGGAGGAGGAGCCATCAAGCAGCTACGCAAACTGCGGGCTCAGATGAAGGACTT 4970  
QY 253 -ThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSe 272  
Db 4971 TCAAGAGAGCTGGAAGATGCCGCTGCTCCAGAGATGAGATCTTTGCCACA---GCCAA 5027  
QY 272 rSerAsnGluProSerLeuLysGluLeuGluLysAspProValTrpLysLysLeuAsnAl 292  
Db 5028 AGAGATGAGAGAGAGAGAGAGCTTGGAGAGAGCTTCATCAGTACAGAGGAGCT 5087

Qy 292 aValLysAenGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArgGlyLe 312  
 Db 5088 CGCCGCCCTGAGAG-----GTCGCAACAACGCGGA 5120  
 Qy 312 uileSerSerGluGluMetAlaLysGluLeuVal---GluLeuSerLysLysAspSerLys 331  
 Db 5121 CCTCGAAGAAGGAGAACTGGCAGAGGAGCTGGCCAGTAGCCTGTGCGGAAGGACGCAC 5180  
 Qy 331 slyshApheLys 335  
 Db 5181 CCAGGACGAGAAG 5193

RESULT 20  
 CO573945  
 LOCUS CO573945 804 bp mRNA linear EST 19-JUL-2004  
 DEFINITION AGENCOURT\_28537271 NIH\_MGC\_246 Rattus norvegicus cDNA clone  
 IMAGE:7372914 5', mRNA sequence.

ACCESSION CO573945  
 VERSION CO573945.1 GI:50386574  
 KEYWORDS EST.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 804)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Daniela S. Gerhardt, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: c9apbs-remail.nih.gov  
 Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical  
 College of Wisconsin  
 cDNA Library Preparation: Open Biosystems  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM15519 row: f column: 16  
 High quality sequence stop: 541.  
 Location/Qualifiers

FEATURES  
 source  
 1..804  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10116"  
 /clone="IMAGE:7372914"  
 /lab\_host="DH10B Tona"  
 /clone\_lib="NIH\_MGC\_246"  
 /note="Organ: liver; Vector: pExpress-1; Site: 1: EcoRV;  
 Site 2: NotI; RNA obtained from testis tissue of 8 wk old  
 animal. Tissues were snap-frozen and kept at -80C before  
 RNA extraction and purification (Tri-reagent method). cDNA  
 was primed using oligo-dT primer:  
 5'-pGACTAGTCTAGTCGAGCGCGCCCT(T)25-3' and cloned into  
 the EcoRV/NotI sites of pExpress-1. Size-selection >1.25kb  
 resulted in an average insert size of 1.7 kb. This a  
 primary library (normalized library is NIH MGC 247) and  
 was constructed by Open Biosystems. Note: This is a  
 NIH\_MGC library"

## ORIGIN

Alignment Scores:  
 Pred. No.: 0.0251 Length: 804  
 Score: 116.50 Matches: 67  
 Percent Similarity: 37.3% Conservative: 38  
 Best Local Similarity: 23.84% Mismatches: 79  
 Query Match: 6.89% Indels: 97  
 DB: 7 Gaps: 14

US-10-724-972A-6352 (1-335) x CO573945 (1-804)  
 Qy 6 ArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeu----- 20  
 Db 60 CGGGAAGCAAGATACAGACAAGAGTCGGGTGTGTGATGTGCTTGTACTCGTGTAAAGC 119  
 Qy 21 -----IleAlaThrAlaAlaCysGlyAsnAsnSerSerSerAsnSerSerLysGlu 37  
 Db 120 AACTGTCCGATTGCTGTAAACACACTTCTTCACAAAC---TCAGCCAATGTTCCATTCTT 176  
 Qy 38 SerSerLysAspGlyValGluIleLysHisGluGlyThrThrLysValProLysHis 57  
 Db 177 ACAGGACAGATTGCGAGAAAATCTCGAGAGAAGAG----- 212  
 Qy 58 ProLysArgValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspVal 77  
 Db 213 -----CAGTTGGTCCAAGGCTTATGTGCCCTT----- 239  
 Qy 78 LysProValGlyIleAla-----AspAspAsnLys-----LysAsn 89  
 Db 240 ---CTTTGGGCAATTCAAATTTATTTCAACGACAACCTCAGTAAAGAACTACACGAAGAG 296  
 Qy 90 ArgIleIleLysProLeuArgAspLysIleGlyLys----- 101  
 Db 297 AAATAAGCAACTAATAGAGAGAGGATTGGCAAGAGAAATTACATAGAGAAACTTGGGA 356  
 Qy 102 -----TyrThrSerValGlyThrArgLysGlnProAsnLeuGlu 114  
 Db 357 TTTATTAGCAAGCATGAGTTATATTTCCAGAGCATCTCAGAAACCCCACTTCCG 416  
 Qy 115 GluIleSerLysLeuLysProAspLeuIleAlaAspAsnAsnArgHisLysGlyIle 134  
 Db 417 -----AGTCCAGATAATATTTGATCATGATGATTTACAAA---CTG 458  
 Qy 135 TyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGlyAspTyr 154  
 Db 459 GTGAAGAAGAACTTGAAGGTGTTATTACCAAGGCTATTATTAAGTCCAGCGAAGAAGAT--- 515  
 Qy 155 AsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGluGly 174  
 Db 516 -----AAGAAAGAGGAGGAAGGT----- 533  
 Qy 175 LysLysArgLeuGluGluHisAspLys----- 183  
 Db 534 AAGAAACCTTAGAACACGATGATAATATTGTCACACTATAAGAATATGATCGGTGAG 593  
 Qy 184 -----LysIleGluGluTyrLysLysGluIleThrMetAspLys-----AsnGln 198  
 Db 594 NCAGACCTTACAACCTGGAGGAACATAAACAGACAGTGTCCACACTGAAATGTGAGAACGAG 653  
 Qy 199 LysValLeuProAlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyr 218  
 Db 654 CAGCTGNCACAGCAGCATCCCCAGCAGCTTCAGATTCAGCAGCAGCAGATCAGTACA 713  
 Qy 219 ValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLys 238  
 Db 714 ACCTCTCAA-----AGTCAGCTCGGG-----AAAGACATCATCATCA 749  
 Qy 239 Gly 239  
 Db 750 GGT 752

## RESULT 21

BJ714899  
 LOCUS BJ714899 832 bp mRNA linear EST 08-MAR-2004  
 DEFINITION BJ714899 MF01FFA cDNA Oryzias latipes cdna clone MF01FFA03302 5',  
 mRNA sequence.  
 ACCESSION BJ714899  
 VERSION BJ714899.1 GI:45255740  
 KEYWORDS EST.  
 SOURCE Oryzias latipes (Japanese medaka)  
 ORGANISM Oryzias latipes





```

/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Eb013122"
/tissue_type="floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_lib="GR_Eb"
/notes="vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; Library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Alignment Scores:
Pred. No.: 0.0292 Length: 816
Score: 116.00 Matches: 68
Percent Similarity: 37.89% Conservative: 40
Best Local Similarity: 23.86% Mismatches: 119
Query Match: 6.86% Indels: 58
DB: 7 Gaps: 11

US-10-724-972A-6352 (1-335) x C0113362 (1-816)

Qy 84 AspaAspAenLysLysAsnArgIleLysProLeuArgAspLysIleGlyLysTyrThr 103
Db 10 GAAGAGGAAAGAAAGAA-----AAGCCG---AAGACAAAGAGGCAAGGAGAAG 57
Qy 104 SerValGlyThrArgLysGlnProAsnLeuGluLeuSerLysLeu----- 119
Db 58 AAGAAAGAAAGAAAGAAAGGATGAGTAGAAGAGCTGAAGAATTAGAAGGGGGAAG 117
Qy 120 -----LysProAspLeuIleAlaAspAsnAsnArgHisLysGlyIleTyrLys 136
Db 118 AAGAAAGAAAGAAAGAAAGGCGGCGAGGAGATGAAGAAAGAA-----AAG 168
Qy 137 AspLeuAsnLysIleAlaProThrIleGluLeu----- 147
Db 169 AAGAAGGATGAGGTAGAAGAAAGATGAAGAATTAGAAGAGGGAAGAAAGAAAGAA 228
Qy 148 -----LysSerPheAspGlyAspTyrAsn 155
Db 229 GACAAAGAGCGGAGGAGAAAGAAAGAAAGAAAGAGGAGGAGGAGGAGGATGAAGAT 288
Qy 156 GluAsnIleAspAlaPheLysThrIleSerLys-----AlaLeuGlyLysGlu 171
Db 289 GAAGATTAGAAGGCGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGGCAAGGAGAGAAG 348
Qy 172 GluGluGlyLysLysArgLeuGlu-----GluHisAspLysLysIleGluTyrLysLys 190
Db 349 AAAGAAAGAAAGAAAGAGGATGAGGTAGATGAAGATGAAGAATTAGAAGCGGAAAGAA 408
Qy 191 GluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeu 210
Db 409 GAAAGAGAAAGAAAGCAAGAGCGGCGAGGAGAAAGAAAGAAAGAAAGAGGATGAG 468
Qy 211 LeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGlu 230
Db 469 GTAGAGAACTCAAGAA-----TTAGAGAGGGAATTAAGAA 507
Qy 231 AlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGln 250
Db 508 AAGAAGAAACACAAAGAGCGGAGGAGAGAAAGAAAGAAAGAAAGAGGATGAGGTA 567
Qy 251 MetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsnLys 270
Db 568 GAAGAAGCTCAAGAAATTAGAAGAGGGAATTAAGAAAGAAAGAAAGCAAGAGGCAAG 627
Qy 271 AlaSerSerAsnGluProSerLeuLysGluLeuGluLysAspProValTrpLysLysLeu 290
Db 628 GAGAAGAGAAAGAAAGAAAGAGGATGAGGTAGAAGAGAT-----GAAAGAAATTA 678
Qy 291 AsnAlaValLysAsnGlnArgValAspIleLeuAspAspLeuTrpAlaArgSerArg 310

```

```
Db 191 AATAAGAGAAATAAAGAGATTAAATGATGGAAGAAAGAAATTAATAAATCAAAAGAGAAAC 250
Qy 104 SerValGlyThrArgLysGlnProAsnLeuGluGluLeuSerLysLeuLysProAspLeu 123
Db 251 ACTCTTCCAAATTCACCAACCAAGATTAGATGAAAGAAAGAAATGAAG----- 301
Qy 124 IleIleAlaAspAsnAsnArgHisLysGlyLysGlyLysLysLysLysLysLysLysLysLys 143
Db 302 -----GAAATCGTAGA-----ACTATAGTCCA 325
Qy 144 ThrIle-----GluLeuLysSerPheAspGlyAspTyrAsnGluAsnIle 158
Db 326 AGGATGATGGGTTAGGACAGAAATTAACAA-----AATAAGAGAGTAGAGAAAGAAAT 382
Qy 159 AspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGluGly----- 174
Db 383 AGAAGATAGAGAAATTAACAAAGAAATTTAGAGTTAGAACACAGAACAAAGAAATCAACTT 442
Qy 175 LysLysArgLeuGluGluHisAspLysLysLysLysLysLysLysLysLysLysLysLysLys 194
Db 443 AAAGAAAGTTAGAGACAGCAAGAAATCAATTTGAAGAAATGAAGAAAGAAATTAATAA 502
Qy 195 AspLys-----AsnGlnLysValLeuProAlaValAlaLysSer 208
Db 503 GAAAGAGGAGAGTTTGAAGAGAAATTAATGAGAG----- 535
Qy 209 GlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer-----GlnLeu 226
Db 536 -----AATAATATACAAATTAATGAATGAATGAATGAATGAATGAATGAATGAATGA 577
Qy 227 GlyPheLysGluAlaLeuSerAspValThrLysGlyLysLysLysLysLysLysLysLys 246
Db 578 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 610
Qy 247 ProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheIle 266
Db 611 -----CTAAAGAGTTCAATTGAAGAG----- 631
Qy 267 MetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLysLysAspProVal 286
Db 632 CAACATATCAAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 691
Qy 287 ---TrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeu 305
Db 692 GAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 750
Qy 306 TrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysLysLysLysLysLysLys 325
Db 751 TGG-----AGTTTAAAGACACTGAACATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 804
Qy 326 SerLys 327
Db 805 TTGAAA 810

RESULT 25
LOCUS CR650378 1483 bp mRNA linear HTC 18-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR650378
VERSION CR650378.1 GI:51146823
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1483)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
; 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
```

```
(E-mail : seque@genoscope.cns.fr - Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
Location/Qualifiers
source 1. 1483
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="Liver"

ORIGIN
Alignment Scores:
Pred. No.: 0.0675 Length: 1483
Score: 116.00 Matches: 62
Percent Similarity: 41.48% Conservative: 50
Best Local Similarity: 22.96% Mismatches: 88
Query Match: 6.86% Indels: 70
Gaps: 14

US-10-724-972A-6352 (1-335) x CR650378 (1-1483)
Qy 50 GlyThrThrLysValProLysHisProLysArgValValValLeuGluTyrSerPheVal 69
Db 499 GGTTCACATCGCATCCCAAAATCCAGAAA----- 528
Qy 70 AspAlaLeuValAlaLeuAspValLysProValGlyLysLysLysLysLysLysLysLys 89
Db 529 -----CTGCTCCGTCGAGCGTCACTCTCTCTCGGTATTTGAGACG---GCCGGA 576
Qy 90 ArgIleLysProLeuArgAspLysLysLysLysLysLysLysLysLysLysLysLysLys 109
Db 577 GGCCTCATGACGGTCTG-----ATCAAGCGCAACACCATCTCTCCACAGCAG 627
Qy 110 GlnProAsnLeuGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 129
Db 628 ACCCAGACCTTTCACACCTACTCTGACAAACGAGCCGCGTCTCATCCAGTTTACGAG 687
Qy 130 ArgHisLysGlyLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 141
Db 688 GGTGAGCGTGCCATGACCGAGGACAAATCTGTGGGCAAGTTCGAGTTGACGGGCATC 747
Qy 142 -----AlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsn 155
Db 748 CCCCTCTCCCTCTGCGGTCTCCAGATCAGGTG---ACTTTTGATCATCGACGCCAAC 804
Qy 156 -----GluAsnIleAspAlaPheLysThrLysLysLysLysLysLysLysLysLys 173
Db 805 GGCATCATGAACGTGTCTGCC-----GTGACAAAGACACTGGCAAGAAACAAAG 855
Qy 174 -----GlyLysLysArgLeuGluGluHisAsp-----Lys 183
Db 856 ATCACCATTACCAATGACAAAGGTCTCTAGTAGGAGGACATTGACCGCATGCTGGCAG 915
Qy 184 LysIleGluGluTyrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 203
Db 916 GAAGCTGAGAGTACAGGCCGAGACGCTCCAGCGGTGACAG----- 960
Qy 204 ValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeu 223
Db 961 GTGTGGCCAAAGATTTGAGTCTGCTGCTTCAAC----- 999
Qy 224 SerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGly---LeuSerLys 242
Db 1000 -----ATGAGTCAACCTGGAGACGAGAACTCGCTGGGAAAGATCATGTGAC 1047
Qy 243 TyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLysSerGlnValAsnProGlu 262
Db 1048 GATGACAGCAGAGAGATTTTGGACAAAGTGAACGAGGTGTCATCAGTGTCTGACAGAAAT 1107
Qy 263 ArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGlu 282
Db 1108 CAG-----ACTGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1155
```



```

ORIGIN
embryo"

Alignment Scores:
Pred. No.: 0.0259 Length: 684
Score: 115.50 Matches: 65
Percent Similarity: 41.04% Conservative: 38
Best Local Similarity: 25.90% Mismatches: 91
Query Match: 6.83% Indels: 57
DB: 5 Gaps: 14

US-10-724-972A-6352 (1-335) x BW253010 (1-684)
QY 69 ValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsnLysLys 88
DB 8 GTTCAGGATCTCTCTGTGGATGTCACCTTTATCTCTGTGATTGAGACC---GCT 64

QY 89 AsnArgIleLeuLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArg 108
DB 65 GGAGGAGTGATGACAGCTTTG-----ATCAAGAGAAACACGACCGTCCCAACCAAA 115

QY 109 LysGlnProAsnLeuGluIleSerLysLeuLysProAspLeuIleAlaAspAsn 128
DB 116 CAATCCCAAAATATTCACCACTTACTCGGACAAACACGAGGTGTGTGATCCAAAGTGATC 175

QY 129 AsnArgHisLeuGlyIleTyrLysAspLeuAsnLysIle----- 141
DB 176 GAGGTGTAACGTACGATGACCAAGGACAAACATATCTTGTAAAGTTCGAGCTCTCCGGC 235

QY 142 -----AlaProThrIleGluLeuLysSerPheAspGlyAspTyr 154
DB 236 ATTCCTCCCTGCACCTCGTGGAGTTCACAGATTGAGGTG---ACGTTGATATCGATGCT 292

QY 155 AsnGlu-----AsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGlu 172
DB 293 AACGGTATCTGAATGTATCTGCT-----GTTGACAAGGACACTGGAAGGAGAAAT 343

QY 173 Glu-----GlyLysLysArgLeuGluGluHisAsp----- 182
DB 344 AAGATCAACATCAACCAACGACAAAGGTGCTCTAGCAAGGAGATATTGACGGGATGGTG 403

QY 183 LysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuPro 202
DB 404 AACGAAGCGGAGAAATATAAGGAAGAA-----GATGAGAAGCAGAAG-----GAG 448

QY 203 AlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe 222
DB 449 AAGATTCAAGCGAAGAACCGACTCGAATCTTATGCGTTCAAC----- 490

QY 223 LeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGly---LeuSer 241
DB 491 -----CTTAATCAACAGTGGAGGATGACAAAGTGAAGATATAAATTTC 535

QY 242 LysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnPro 261
DB 536 GACGAAGATATAATCAGCGATTCTTAACAAAGCTAAAGAAAGTTTTCGACTGGTGGAAAT 595

QY 262 GluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeu 281
DB 596 AACCCAG-----ACTCCGAAAAGGACGAATATGAATTCACAAAAAAGAACTT 643

QY 282 GluLys-----AspProValTrpLysLysLeu 290
DB 644 GAGAAATTTGCCAACCTTATTATGACAAAACTTT 676

RESULT 28
LOCUS BW254793
DEFINITION BW254793 Nori Satoh unpublished cDNA library, tailbud embryo Ciona
intestinalis cDNA clone citb073p02 5', mRNA sequence.
ACCESSION BW254793
VERSION BW254793.1 GI:24834711
KEYWORDS EST.

SOURCE
ORGANISM
Ciona intestinalis
Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cionidae; Ciona.
REFERENCE
1 (bases 1 to 690)
Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
Expressed genes in Ciona intestinalis (2002c)
Unpublished (2002)
Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@scid.ian.zool.kyoto-u.ac.jp.
FEATURES
Location/Qualifiers
1..690
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="citb073p02"
/tissue_type="whole animal"
/dev_stage="tailbud embryo"
/clone_lib="Nori Satoh unpublished cDNA library, tailbud
embryo"

ORIGIN
Alignment Scores:
Pred. No.: 0.0263 Length: 690
Score: 115.50 Matches: 65
Percent Similarity: 41.04% Conservative: 38
Best Local Similarity: 25.90% Mismatches: 91
Query Match: 6.83% Indels: 57
DB: 5 Gaps: 14

US-10-724-972A-6352 (1-335) x BW254793 (1-690)
QY 69 ValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsnLysLys 88
DB 19 GTTCAGGATCTCTCTGTGGATGTCACCTTTATCTCTGTGATTGAGACC---GCT 75

QY 89 AsnArgIleLeuLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArg 108
DB 76 GGAGGAGTGATGACAGCTTTG-----ATCAAGAGAAACACGACCGTCCCAACCAAA 126

QY 109 LysGlnProAsnLeuGluIleSerLysLeuLysProAspLeuIleAlaAspAsn 128
DB 127 CAATCCCAAAATATTCACCACTTACTCGGACAAACACGAGGTGTGTGATCCAAAGTGATC 186

QY 129 AsnArgHisLeuGlyIleTyrLysAspLeuAsnLysIle----- 141
DB 187 GAGGTGTAACGTACGATGACCAAGGACAAACATATCTTGTAAAGTTCGAGCTCTCCGGC 246

QY 142 -----AlaProThrIleGluLeuLysSerPheAspGlyAspTyr 154
DB 247 ATCCCTCCCTGCACCTCGTGGAGTTCACAGATTGAGGTG---ACGTTGATATCGATGCT 303

QY 155 AsnGlu-----AsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGlu 172
DB 304 AACGGTATCTGAATGTATCTGCT-----GTTGACAAGGACACTGGAAGGAGAAAT 354

QY 173 Glu-----GlyLysLysArgLeuGluGluHisAsp----- 182
DB 355 AAGATCAACATCAACCAACGACAAAGGTGCTTACCAAGAAAGATATTGACGGGATGGTG 414

QY 183 LysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuPro 202
DB 415 AACGAGGCGGAGAAATATAAGGAAGAA-----GATGAGAAGCAGAAG-----GAG 459

QY 203 AlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe 222
DB 460 AAGATTCAAGCGAAGAACCGACTCGAATCTTATGCGTTCAAC----- 501

```



```
QY 223 LeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGly---LeuSer 241
Db 502 -----CTTAAATCAACGGTGGAGGATGACAAAGTGAAGATTAAGATTCA 546
QY 242 LysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnPro 261
Db 547 GACGAAGATAAGTCAGCGATTCTTAACAAGCTAAAGAAGTTTGGACTGGTTGGAAT 606
QY 262 GluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeu 281
Db 607 AACACAG-----ACTCGGAAAGAGGACGAATATGAATTCACACAAAGAACTT 654
QY 282 GluLys-----AspProValTrpLysLysLeu 290
Db 655 GAGAAAAATTGCCAACCTATTATGACTAAACTT 687

RESULT 29
BW222839
LOCUS BW222839 692 bp mRNA linear EST 06-NOV-2002
DEFINITION BW222839 Nori Satoh unpublished cDNA library, larva Ciona
intestinalis cDNA clone cili051j23 5', mRNA sequence.
ACCESSION BW222839
VERSION BW222839.1 GI:24742218
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
REFERENCE 1 (bases 1 to 692)
AUTHORS Satou,Y., Shin-i,T., Kohara,Y. and Satoh,N.
TITLE Expressed genes in Ciona intestinalis (2002c)
JOURNAL Unpublished (2002)
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoheascidian.zool.kyoto-u.ac.jp.

FEATURES
source
location/Qualifiers
1..692
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="cili051j23"
/tissue_type="whole animal"
/dev_stage="larva"
/clone_lib="Nori Satoh unpublished cDNA library, larva"

ORIGIN
Alignment Scores:
Pred. No.: 0.0264 Length: 692
Score: 115.50 Matches: 65
Percent Similarity: 41.04% Conservative: 38
Best Local Similarity: 25.90% Mismatches: 91
Query Match: 6.83% Indels: 57
DB: 5 Gaps: 14

US-10-724-972A-6352 (1-335) x BW222839 (1-692)
QY 69 ValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsnLysLys 88
Db 21 GTTCAGGATCTTCTCTGTGGATGTGACCTTTATCTCTGGTATTGAGACC---GCT 77
QY 89 AsnArgIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArg 108
Db 78 GGAGGAGTGATGACAGCTTTG-----ATCAAGAGAAACACACCGTCCCAACCAA 128
QY 109 LysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsn 128
Db 129 CAATCCCAAAATATTCACCACTTACTCGGACCAACCAACGAGGTGTGTATCCAGTGATC 188
QY 129 AsnArgHisLysGlyIleTyrLysAspLeuAsnLysIle----- 141
```

```
Db 189 GAGGGTGAACGTACGATGACCAAGGACAAACATCTACTTGGTAAAGTTCGAGCTCTCCGCG 248
QY 142 -----AlaProThrIleGluLeuLysSerPheAspGlyAspTyr 154
Db 249 ATTCCCTCGCCTCGTGGAGATCCACAGATTGAGGTG---ACGTTGCATATCGATGCT 305
QY 155 AsnGlu-----AsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 172
Db 306 AACGGTATCTTTGAATGTATCTGCT-----GTTGACAAGAGCACTGGAAGGAGAAT 356
QY 173 Glu-----GlyLysLysArgLeuGluGluHisAsp----- 182
Db 357 AAGATCACAATCACCAACGACAAAGGTGCTTTAGCAAGGAGATATTGAGCGGATGGTG 416
QY 183 LysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuPro 202
Db 417 AACGNAAGCGGAGAAATATAAGGAAGAA-----GATGAGACGACAGAG-----GAG 461
QY 203 AlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe 222
Db 462 AAGATTCAAGCGAAGAACGACTCGAATCTTTATGCGTTCAAC----- 503
QY 223 LeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGly---LeuSer 241
Db 504 -----CTTAAATCAACAGTCAGGAGTGCACAAAGTCAAAAGATAAAATTTCA 548
QY 242 LysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnPro 261
Db 549 GACGAGATAAATCAGCGATTCTTAACAAAGACTAAAGAGTTTGGACTGGTTGGAAT 608
QY 262 GluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeu 281
Db 609 AACACAG-----ACTCGGAAAGAGGACGAATATGAATTCACACAAAGAACTT 656
QY 282 GluLys-----AspProValTrpLysLysLeu 290
Db 657 GAGAAAAATTGCCAACCTATTATGACAAACTT 689

RESULT 30
BW235942
LOCUS BW235942 700 bp mRNA linear EST 07-NOV-2002
DEFINITION BW235942 Nori Satoh unpublished cDNA library, tailbud embryo Ciona
intestinalis cDNA clone citb052c24 5', mRNA sequence.
ACCESSION BW235942
VERSION BW235942.1 GI:24758131
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
REFERENCE 1 (bases 1 to 700)
AUTHORS Satou,Y., Shin-i,T., Kohara,Y. and Satoh,N.
TITLE Expressed genes in Ciona intestinalis (2002c)
JOURNAL Unpublished (2002)
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoheascidian.zool.kyoto-u.ac.jp.

FEATURES
source
location/Qualifiers
1..700
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="citb052c24"
/tissue_type="whole animal"
/dev_stage="tailbud embryo"
/clone_lib="Nori Satoh unpublished cDNA library, tailbud embryo"

ORIGIN
```

```

Alignment Scores:
Pred. No.: 0.0268 Length: 700
Score: 115.50 Matches: 65
Percent Similarity: 41.04% Conservative: 38
Best Local Similarity: 25.90% Mismatches: 91
Query Match: 6.83% Indels: 57
DB: 5 Gaps: 14

US-10-724-972A-6352 (1-335) x BW235942 (1-700)
Qy 69 ValAspAlaLeuValAlaLeuAspValLysProValGlyLeuAlaAspAsnLysLys 88
Db 26 GTTCAGGATCTCTCTGTTGGATGTCACCTTTATCTCTGTTGATGAGACC---GCT 82
Qy 89 AsnArgIleLeuLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArg 108
Db 83 GGAGGAGTGATGACAGCTTTG-----ATCAAGAGAGAAACACCGCTCCCAACCAAA 133
Qy 109 LysGlnProAsnLeuGluLysLeuSerLysLysLysProAspLeuIleAlaAspAsn 128
Db 134 CAATCCCAAAATATTCACCACTTACCGACCAACCAAGGCTGTTGATCAAGTGATC 193
Qy 129 AsnArgHieLysGlyIleTyrLysAspLeuAsnLysIle----- 141
Db 194 GAGGAGTGAACGTACGATACCAAGAGCAACACATCTTGTGTAAGTTCGAGCTCTCCGGC 253
Qy 142 -----AlaProThrIleGluLeuLysSerPheAspGlyAspTyr 154
Db 254 ATCCCCCTGCACCTCGTGGAGTTCACAGATTTGAGTG---ACGTTGATATCGATGCT 310
Qy 155 AsnGlu-----AsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGlu 172
Db 311 AACGGTATCTTGAATGTATCTGCT-----GTTGACAGAGCACTGGAAGGAGAA 361
Qy 173 Glu-----GlyLysLysArgLeuGluGluHieAsp----- 182
Db 362 AAGATCAATCAATCAACAGCAAGAGTCTTACCAAGGAGATATTGAGCGGATGTG 421
Qy 183 LysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuPro 202
Db 422 AACGAGCGCGAGAAATATAAGGAGAA-----GATGAGACAGCAAG-----GAG 466
Qy 203 AlaValAlaAlaLysSerGlyLysLeuAlaHieProSerAsnSerTyrValGlyGlnPhe 222
Db 467 AAGATTCAAGCGCAAGACGAGCTCGAATCTTATGCGTTCAAC----- 508
Qy 223 LeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGly---LeuSer 241
Db 509 -----CTTAATCAACGCTGGAGGATGCAAAAGTCAAGATATAAATTTCA 553
Qy 242 LysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnPro 261
Db 554 GACGAAGATATAATCAGCGATTCTTAACAAAGCTAAAGAAAGTTTGGAGCTGTTGAAAAT 613
Qy 262 GluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeu 281
Db 614 AACCAAG-----ACTCCGAAAGAGGAGCAATATGAATTCACAAAGAAAGAACTT 661
Qy 282 GluLys-----AspProValTrpLysLysLeu 290
Db 662 GAGAAATTTGCCAACCTTATTATGACAAAATTT 694

RESULT 31
LOCUS BW248137
DEFINITION BW248137 Nori Satoh unpublished cDNA library, tailbud embryo Ciona
intestinalis cDNA clone citb076005 5', mRNA sequence.
ACCESSION BW248137
VERSION BW248137.1
KEYWORDS GI:24828055
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis

```

Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;  
Phlebobranchia; Clonidae; Ciona.

REFERENCE  
AUTHORS 1 (bases 1 to 702)  
TITLE Satou,Y., Shin-i,T., Kohara,Y. and Satoh,N.  
JOURNAL Expressed genes in Ciona intestinalis (2002c)  
COMMENT Unpublished (2002)  
CONTACT: Nori Satoh  
DEPARTMENT: Department of Zoology  
INSTITUTION: Kyoto University  
ADDRESS: Sakyo-ku, Kyoto, Kyoto 606-8502, Japan  
TELEPHONE: Tel: 81-75-753-4081  
FAX: Fax: 81-75-705-1113  
EMAIL: Email: satoh@cicdian.zool.kyoto-u.ac.jp.

FEATURES  
source  
1..702  
/organism="Ciona intestinalis"  
/mol\_type="mRNA"  
/db\_xref="taxon:7719"  
/clone="citb076005"  
/tissue\_type="whole animal"  
/dev\_stage="tailbud embryo"  
/clone\_lib="Nori Satoh unpublished cDNA library, tailbud embryo"

## ORIGIN

Alignment Scores:  
Pred. No.: 0.0269 Length: 702  
Score: 115.50 Matches: 65  
Percent Similarity: 41.04% Conservative: 38  
Best Local Similarity: 25.90% Mismatches: 91  
Query Match: 6.83% Indels: 57  
DB: 5 Gaps: 14

US-10-724-972A-6352 (1-335) x BW248137 (1-702)

```

Qy 69 ValAspAlaLeuValAlaLeuAspValLysProValGlyLeuAlaAspAsnLysLys 88
Db 11 GTTCAGGATCTCTCTGTTGGATGTCACCTTTATCTCTGTTGATGAGACC---GCT 67
Qy 89 AsnArgIleLeuLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArg 108
Db 68 GGAGGAGTGATGACAGCTTTG-----ATCAAGAGAGAAACACGCGTCCCAACCAAA 118
Qy 109 LysGlnProAsnLeuGluLysLeuSerLysLysLysProAspLeuIleAlaAspAsn 128
Db 119 CAATCCCAAAATATTCACCACTTACCGACCAACCAACGAGTGTGTTGATCCAAAGTGATC 178
Qy 129 AsnArgHieLysGlyIleTyrLysAspLeuAsnLysIle----- 141
Db 179 GAGGAGTGAACGTACGATGACCAAGAGCAACACATCTTGTGTAAGTTCGAGCTCTCCGGC 238
Qy 142 -----AlaProThrIleGluLeuLysSerPheAspGlyAspTyr 154
Db 239 ATTCCTCCCTGCACCTCGTGGAGTTCACAGATTTGAGTG---ACGTTGATATCGATGCT 295
Qy 155 AsnGlu-----AsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGlu 172
Db 236 AACGGTATCTTGAATGTATCTGCT-----GTTGACAGAGCACTGGAAGGAGAAAT 346
Qy 173 Glu-----GlyLysLysArgLeuGluGluHieAsp----- 182
Db 347 AAGATCACAATCAACCAAGCAAAAGTCTTACCAAGGAGATATTGAGCGGATGTG 406
Qy 183 LysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuPro 202
Db 407 AACGAAGCGCGAGAAATATAAGGAGAA-----GATGAGACAGCAAG-----GAG 451
Qy 203 AlaValAlaAlaLysSerGlyLysLeuAlaHieProSerAsnSerTyrValGlyGlnPhe 222
Db 452 AAGATTCAAGCGAGAGAGAGCTCGAATCTTATGCGTTCAAC----- 493
Qy 223 LeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGly---LeuSer 241
Db 493 -----

```

```

Db      494  -----CTTAATCAACAGTGGAGGATGACAAAGTGAAGATAAAATTTCA 538
Qy      242  LysTyLeuLysGlyProTyLeuGlnMetAsnThrGluThrLeuSerGlnValAsnPro 261
Db      539  GACGAAGATAAATCAGCATTTTAAACAAAGCTAAAGAGTTTGGACTGGTTGGAAAT 598
Qy      262  GluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeu 281
Db      599  AACGAG-----ACTGCCGAAAGAGGACGAATATGAATTCACAAAAAGAACTT 646
Qy      282  GluLys-----AspProValTrpLysLysLeu 290
Db      647  GAGAAATTCGCAACCTATTATGACAAACTT 679

RESULT 32
BW441129
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BW441129 713 bp mRNA linear EST 09-JUN-2004
BW441129 Nori Satoh unpublished cDNA library, juvenile whole animal
Ciona intestinalis cDNA clone cijv044124 5', mRNA sequence.
BW441129
BW441129.1 GI:48494172
EST.
Ciona intestinalis
Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Ciona.
1 (bases 1 to 713)
Satou,Y., Nakayama,A., Shin-i,T., Kohara,Y. and Satoh,N.
Expressed genes in Ciona intestinalis (2004b)
Unpublished (2004)
Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source
1. .713
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="cijv044124"
/tissue_type="whole animal"
/dev_stage="juvenile"
/clone_lib="Nori Satoh unpublished cDNA library, juvenile
whole animal"

ORIGIN
Alignment Scores:
Pred. No.: 0.0275 Length: 713
Score: 115.50 Matches: 65
Percent Similarity: 41.04% Conservative: 38
Best Local Similarity: 25.90% Mismatches: 91
Query Match: 6.83% Indels: 57
DB: 5 Gaps: 14

US-10-724-972A-6352 (1-335) x BW441129 (1-713)

Qy      69  ValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLysLys 88
Db      34  GTTCAGGATCTTCTCTGTGGATGTGGACCTTTATCTCTGTATTGAGACC---GCT 90
Qy      89  AsnArgIleLysProLeuArgAspLysIleGlyLysTyThrSerValGlyThrArg 108
Db      91  GGAGGAGTGATGACGCTTG-----ATCAAGAGAAACACCGCTCCCAACCAA 141
Qy      109  LysGlnProAsnLeuGluIleLysSerLysLeuLysProAspLeuIleAlaAspAsn 128
Db      142  CAATCCCAATATTCACCACTTACTCGACAAACCAACCGAGTGTGTGATCCCAAGTG 201
Qy      129  AsnArgHisGlyLysLeuLysAspLeuLysLysIle----- 141

```

```

Db      202  GAGGGTGAACGTACGATGACCAAGGACAAACATACCTTGGTTAAGTTCGAGCTCTCCGCG 261
Qy      142  -----AlaProThrIleGluLeuLysSerPheAspGlyAspTy 154
Db      262  ATCCCCCTGCACCTCGTGGAGTTCACAGATTTGAGGTG---ACGTTGATATCGATGCT 318
Qy      155  AsnGlu-----AsnIleAspAlaPheLysThrIleSerLysAlaLeuLysGluGlu 172
Db      319  AACGGTATCTTGAATGTATCTGCT-----GTTGACAAGAGCACTGGAAGGAGAAT 369
Qy      173  Glu-----GlyLysLysArgLeuGluGluHisAsp----- 182
Db      370  AAGATCACAATCACCACCAAGCAAGAGTCTCTTAGCAAGAGATATTGAGCGGATGGTG 429
Qy      183  LysLysIleGluGluTyTrpLysLysGluIleThrMetAspLysAsnGlnLysValLeuPro 202
Db      430  AACGAGGCGCGAGAAATATAAGGAAGAA-----GATGAGAAGCAGAAG---GAG 474
Qy      203  AlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyValGlyGlnPhe 222
Db      475  AAGATTCAAGCGAAGAACGACTCGAATCTTATGGCTTCAAC----- 516
Qy      223  LeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGly---LeuSer 241
Db      517  -----CTTAATCAACGGTGGAGGATGACAAAGTGAAGATAAAATTTCA 561
Qy      242  LysTyLeuLysGlyProTyLeuGlnMetAsnThrGluThrLeuSerGlnValAsnPro 261
Db      562  GACGAAGATAAATCAGCGATTCTTAACAAAGCTAAAGAGTGTGGACTGGTTGGAAT 621
Qy      262  GluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeu 281
Db      622  AACCCAG-----ACTCCGAAAGAGGACGAATATGAATTCACAAAAAGAACTT 669
Qy      282  GluLys-----AspProValTrpLysLysLeu 290
Db      670  GAGAAATTCGCAACCTATTATGACAAACTT 702

RESULT 33
BW055754
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BW055754 753 bp mRNA linear EST 19-OCT-2002
BW055754 Nori Satoh unpublished cDNA library, blood cells Ciona
intestinalis cDNA clone cibd084920 5', mRNA sequence.
BW055754
BW055754.1 GI:24156450
EST.
Ciona intestinalis
Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Ciona.
1 (bases 1 to 753)
Satou,Y., Satake,M., Azumi,K., Nonaka,M., Shin-i,T., Kohara,Y. and
Satoh,N.
Expressed genes in Ciona intestinalis (2002)
Unpublished (2002)
Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source
1. .753
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="cibd084920"
/tissue_type="blood cells"
/clone_lib="Nori Satoh unpublished cDNA library, blood
cells"

ORIGIN

```

Alignment Scores:		0.0297	Length:	753
Pred. No.:	115.50	Matches:	65	
Score:	41.04%	Conservative:	38	
Percent Similarity:	25.90%	Mismatches:	91	
Best Local Similarity:	6.83%	Indels:	57	
Query Match:	5	Gaps:	14	
DB:				
US-10-724-972A-6352 (1-335) x BW055754 (1-753)				
Qy	69	ValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsnLysLys	88	
Db	81	GTTCAGGATCTTCTCTGTGGATTCGACCTTTATCTCTGTGGTATTGAGACC---	137	
Qy	89	AsnArgIleLeuLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArg	108	
Db	138	GGAGGAGTGATGACAGCTTTG-----ATCAAGAGAAACACCCCGTCCCAACCAAA	188	
Qy	109	LysGlnProAsnLeuGluGluIleSerLysLysProAspLeuIleAlaAspAsn	128	
Db	189	CAATCCCAATATTCACCATCTTCTCGGACCAACACAGGTGTGTGTATCCAGTGTAC	248	
Qy	129	AsnArgHisLysGlyIleTyrLysAspLeuAsnLysIle-----	141	
Db	249	GAGGCGTGAACGTACGATGACCAAGACAAACATCTTGTGTAAGTTCGAGCTCTCCGGC	308	
Qy	142	-----AlaProThrIleGluLeuLysSerPheAspGlyAspTyr	154	
Db	309	ATCCCCCTGCACCTCGTGGATTCACAGATTGAGGTG---ACGTTCGATATCGATGCT	365	
Qy	155	AsnGlu-----AsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGlu	172	
Db	366	AACGCTATCTGAATGTATCTGCT-----GTTGACAGACACTGGAAGAGGAAT	416	
Qy	173	Glu-----GlyLysLysArgLeuGluGluHisAsp-----	182	
Db	417	AAGATCACAATCACCAACGACAAAGTCTCTTAGCAAGGAAGATATTGACGGAATGTTG	476	
Qy	183	LysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuPro	202	
Db	477	AACGAGCGGAGAAATATAAGNAGAA-----GATGAGAAGCAAG-----GAG	521	
Qy	203	AlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe	222	
Db	522	AAGATTCAGCGAAGACGGTCTCGAATCTTATGCGTTCAAC-----	563	
Qy	223	LeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGly---	241	
Db	564	-----CTTAAATCAACGGTGAAGATGACAAAGTGAAGATAAAATTTCA	608	
Qy	242	LysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnPro	261	
Db	609	GACGAAGATAAATACGCGATTCTTAACAAAGCTAAAGAAAGTTTGGATGTTGGAAT	668	
Qy	262	GluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeu	281	
Db	669	AAACAG-----ACTCCCAAAAGGAGCAATATGAATTCACCAACAAAGAACTT	716	
Qy	282	GluLys-----AspProValTrpLysLysLeu	290	
Db	717	GAGAAATTTGCCAACCTTATTATGACAAACTT	749	
RESULT 34				
BI905625				
LOCUS				
DEFINITION				
603167626F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5255742 5',				
mRNA sequence.				
ACCESSION				
BI905625				
VERSION				
BI905625.1 GI:16168168				
KEYWORDS				
EST.				
SOURCE				
Mus musculus (house mouse)				
ORGANISM				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				

REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 981)
TITLE	NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999)
	Contact: Robert Strausberg, Ph.D.
	Email: cgabs-r@mail.nih.gov
	Tissue Procurement: Gilbert Smith, Ph.D.
	cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
	DNA Sequencing by: Incyte Genomics, Inc.
	Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
	http://image.llnl.gov
	Plate: LLAM1646 row: c column: 07
	High quality sequence start: 51
	High quality sequence stop: 950.
FEATURES	Location/Qualifiers
source	1..981
	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="CZECH II"
	/db_xref="taxon:10090"
	/clone="IMAGE:5255742"
	/tissue_type="pooled lung tumors"
	/lab_host="DH10B (phage-resistant)"
	/clone_lib="NCI CGAP Lu33"
	/note="Organ: lung; Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCTCTGTTTCTTTTCTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "
ORIGIN	
Alignment Scores:	
Pred. No.:	0.043
Score:	Length: 981
Percent Similarity:	Matches: 61
Best Local Similarity:	Conservative: 40
Query Match:	Mismatches: 106
DB:	Indels: 55
	Gaps: 9
US-10-724-972A-6352 (1-335) x BI905625 (1-981)	
Qy	87 LysLysAsnArgIleLeuLysProLeuArgAspLysIleGlyLysTyrThrSerValGly 106
Db	106 AAGGACAGCAGCTCAGCAGCAGCTGAAGACCTGTCCCAAAAGTACGACGCTGAAG 165
Qy	107 ThrArgLysGlnProAsnLeuGluGluLeuLysLysProAspLeuIleAla 126
Db	166 AGCGAGAGAGAGCTGTAGAGAAAGCCAGCAGCGCTCCGAGATCTCGCAGCC 225
Qy	127 AspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGlu 146
Db	226 CAAAATCTTCTGCAGAAAG-----CAGCGGTCTCCGCTGGAGCAG 264
Qy	147 LeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLys 166
Db	265 GTGGAGGCTCTGAAAAAATCTCTCAACGCGACCATTTGAGCAGTTGAAAG----- 312
Qy	167 AlaLeuGlyLysGluGluGluGlyLysLysArg---LeuGluGluHisAspLysLysIle 185
Db	313 -----GAGGAGCTGAGGAGTAAGCAGAGGTGTCTGGAGCGCGACGACGCGTG 363
Qy	186 GluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAla 205
Db	364 AGCCAGCTTCAGCAGCTGCTGAGAAATCAGAAAGAACTCTCTCGGTGACCTCGCGGAC 423

```
QY 206 AlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyClnPheLeuSerGln 225
Db 424 CTGAAG-----CTAAAGGAAGCCCTCGAAGAAAGAGGTTGGGATCATGAAGCCAGC 474
QY 226 LeuGlyPheGlyAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLys 245
Db 475 CTGAGGAAAGAGGAGAAAGCCAGAGAAACCAAGGAAGTCTCCAAA-----525
QY 246 GlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPhe 265
Db 526 -----CTCCAGACGGAGGTTTCAGACCACCAAGCAG-----555
QY 266 IleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGluLysAspPro 285
Db 556 -----GCCTGAAGAAGTCTTACAGACCCAGAGAG 582
QY 286 Val-----TrpLysLysLeuAsnAlaValLys-----294
Db 583 GTCTGTCGACATGTCGAAGTACAAAGCCACCAAGACGACTTGGAGACCAGATTTCCAAAC 642
QY 295 ----AsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArgGlyLeuLle 313
Db 643 TTAAATGACAAACTGCCAGCCTGACACGGAAGTACGACCAAGCGTGTGAGGAGAGGTC 702
QY 314 SerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAspSerLysLysAsp 333
Db 703 TCTGCCAAGGAC---GAGAAGGAGGCTGCTCCACCTGAGCATCGAGCAGGAGATCAGGGAC 759
QY 334 AsnLys 335
Db 760 CAGAAG 765

RESULT 35
AY420059
LOCUS Mus musculus HCM7097 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY420059
VERSION AY420059.1 GI:39776016
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 4107)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 4107)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence as made by sequencing genomic exons and ordering them
based on alignment.
FEATURES
source Location/Qualifiers
1..4107
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>4107
/locus_tag="HCM7097"
gene
ORIGIN
```

```
Alignment Scores:
Pred. No.: 0.321 Length: 4107
Score: 115.50 Matches: 61
Percent Similarity: 38.55% Conservative: 40
Best Local Similarity: 23.28% Mismatches: 106
Query Match: 6.83% Indels: 55
DB: 9 Gaps: 9

US-10-724-972A-6352 (1-335) x AY420059 (1-4107)

QY 87 LysLysAsnArgIleLeuLysProLeuArgAspLysIleGlyLysTyrThrSerValGly 106
Db 3067 AAGGCAGACGAGCTCAGCAGCAGCTGAAGACCTGTCCCAAAAGTACAGCGACGTGAAG 3126
QY 107 ThrArgLysGlnProAsnLeuGluGluLysSerLysLeuLysProAspLeuLleAla 126
Db 3127 AGCGAGAGAGAGAGCTGTAGAGAAAGGCAAGCAGCGCTCCGAGATCTGGCAGCC 3186
QY 127 AspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGlu 146
Db 3187 CAAATCTTCTGCAGAG-----CAGCCGGTCCCGTGGAGCAG 3225
QY 147 LeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLys 166
Db 3226 GTGGAGGCTCTGAAAAAATCTCTCAACGGCACCATTGACGAGTTGAAG-----3273
QY 167 AlaLeuGlyLysGluGluGluGlyLysLysArg---LeuGluGluHisAspLysLysIle 185
Db 3274 -----GAGGAGCTGAGGAGTAAGCAGAGGTCTCTGGAGCGCGAGCAGCAGCGGTG 3324
QY 186 GluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAla 205
Db 3325 AGCCAGCTTTCAGCAGCTCTGGAGAAATCAGAAGAACTCTCTCGTGACCTCGCGGAGCAG 3384
QY 206 AlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGln 225
Db 3385 CTGAAG-----CTAAGGAAGCCCTCGAAGAAAGAGTTGGATCATGAAGCCAGC 3435
QY 226 LeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLysSerLysTyrLeuLys 245
Db 3436 CTGAGAGAAAGGAAGAGAGAAAGCCAGAAAGAAACCAAGGAAGTCTCCAAA-----3486
QY 246 GlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPhe 265
Db 3487 -----CTCCAGACGAGGTTTCAGACCAACCAAGCAG-----3516
QY 266 IleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLysAspPro 285
Db 3517 -----GGCTGAGAACTTAGAGACCCAGAGAG 3543
QY 286 Val-----TrpLysLysLeuAsnAlaValLys-----294
Db 3544 GTCTGTCGACATGTCGAAGTACAAAGCCACCAAGACGACTTGGAGACCAGATTTCCAAC 3603
QY 295 ----AsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArgGlyLeuLle 313
Db 3604 TTAAATGACAAACTGTGCCAGCCTGAACAGGAAGTACGACCAAGCGCTGTGAGGAGAGGTC 3663
QY 314 SerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAspSerLysLysAsp 333
Db 3664 TCTGCCAAGGAC---GAGAAGGAGTCTCTCCACCTGAGCATCGAGCAGGAGATCAGGGAC 3720
QY 334 AsnLys 335
Db 3721 CAGAAG 3726

RESULT 36
CR426562
LOCUS CR426562
DEFINITION CR426562 XGC-tailbud Xenopus tropicalis cdna clone TTBA033p07 5',
mRNA sequence.
ACCESSION CR426562
VERSION CR426562.1 GI:48919971
```

```

EST.
SOURCE  Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 785)
AUTHORS  Croning, M.D.R., Ashurst, J.L., Taylor, R., Garrett, N. and Rogers, J.
TITLE  Sanger Xenopus tropicalis EST project 2001 (2004)
JOURNAL  Unpublished (2004)
COMMENT  Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TBA033p07.p1kSP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Nigel Garrett.
Seq primer: SP6.
FEATURES             Location/Qualifiers
     source            1..785
     /organism="Xenopus tropicalis"
     /mol_type="mRNA"
     /db_xref="taxon:8364"
     /clone="TBA033p07"
     /dev_stage="tailbud (stage 28-30)"
     /lab_host="Escherichia coli DH10B."
     /clone_lib="XGC-tailbud"
     /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dt primed from Sug of poly A+ RNA from tailbud.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end."
ORIGIN
Alignment Scores:
Pred. No.:          0.0358      Length:      785
Score:              115.00      Matches:    59
Percent Similarity: 37.2%      Conservative: 27
Best Local Similarity: 25.54%  Mismatches: 99
Query Match:        6.80%      Indels:     46
DB:                  7         Gaps:        6

US-10-724-972A-6352 (1-335) x CR426562 (1-785)
Qy 76 AspValLysProValGlyLeuAlaAspAsnLysLysAsnArgIleLeuLysProLeu 95
Db 91 GATGTGAAGAGGTTAAAGTTCTGTGTAATTCGCAAACTTACTCGACACATCGCAAGCTG 150
Qy 96 ArgAspLysIleGlyLysTyrThrSerVal-----GlyThrArgLysGlnProAsn 112
Db 151 AAATCTAAAAGGGGACAGAGAACTAATCTTAAAGATCAGAGGAGCCAGAGATTG 210
Qy 113 LeuGluGluLe-----SerLysLysLysProAspLeuIleIleAlaAspAsn 129
Db 211 CTTGAGGAATCCATCTCTGTTAAGGAGCTCAACACAGATGATGTTACTAAAACTGCTTTA 270
Qy 130 ArgHisLysGlyIleTyr-----Lys 136
Db 271 AGGAAGAAATCAGTGTGTGAAAAAGATGCAAAAAGCTTAACAGCACTGCTGAGGAGCGA 330
Qy 137 AspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGlu 156
Db 331 GCATTAGCAAGACTGGCCACACACACCTTTGCTTAAAGCAAAAATCACAGCTATTAAAGAG 390
Qy 157 AsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLys---GluGluGluGlyLys 175
Db 391 GCTATTAAAGCTTTTAAAGATGCCAAGAAAACCTGCTGCTGAAGGGGAGGGAAGAGAG 450
Qy 176 LysAspGLeuGluGluHisAspLysLysIleGluGluTyrLysLysGluIleThrMetAsp 195
Db 451 AAGGATGAACCTGACAGGTCAAAAGATTAAAGAAACCAAAAACCTGTTACAAACCAAG 510
Qy 196 LysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSer 215

```

---

```

Db 511 CTAATAAAATACAGAAGAGATAAATACAGAAAAGACATGTTAAAGAGAGAAATGC 570
Qy 216 AsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAsp 235
Db 571 AAAAAC-----CTGCTAGAAGAT 588
Qy 236 ValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThr 255
Db 589 TCAGATAAGGCACTGAAAAGCTCTTGAACCTATTATGTGCAAGAGAACTTGCTCTGAG 648
Qy 256 LeuSerGlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSerAsnGlu 275
Db 649 CAGACAGCAGAAAACACAGAA-----CAG 672
Qy 276 ProSerLeuLysGluLeuGluLysAspProVal 286
Db 673 CCGAAGGCCCAAGATGTGGAAGGCCACCTGTG 705

RESULT 37
LOCUS CR731171
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR731171
VERSION CR731171.1 GI:51229441
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1120)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : segre@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES             Location/Qualifiers
     source            1..1120
     /organism="Tetraodon nigroviridis"
     /mol_type="mRNA"
     /db_xref="taxon:99883"
     /tissue_type="fish"
ORIGIN
Alignment Scores:
Pred. No.:          0.059      Length:      1120
Score:              115.00      Matches:    60
Percent Similarity: 40.48%      Conservative: 59
Best Local Similarity: 20.41%  Mismatches: 99
Query Match:        6.80%      Indels:     76
DB:                  3         Gaps:        11

US-10-724-972A-6352 (1-335) x CR731171 (1-1120)
Qy 78 LysProValGlyIleAlaAspAsnLysLys----- 88
Db 162 CGACCGAGCGACACACCGAGCGCGACAGAAAGCGCGCGAGGACAGTGCACACAGCTG 221
Qy 89 -----AsnArgIleIleLysProLeuArgAspLysIleGlyLys 101
Db 222 GAGGAGGAGCTGTGGGTCTCGAAGAAAGCTGAAGGAGTCAAGATGATGCTGACCAAG 281
Qy 102 TyrThrSerValGlyThrArgLysGlnProAsnLeuGluIleSerLysLysPro 121
Db 282 TACTCCGAGTCCCTGAAGGACGCCCGAGAGAGCTGGAGCAGCGGAGGAGGAGGAGGAGCA 341
Qy 122 AspLeu-----IleIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeu 138

```

```
Db 342 GATCGGAGCGGACGTGGCTCTCTGAACCGCGGATCCAGTGGTGAGGAGGCTG 401
Qy 139 AsnLysIleAlaProThrIle-----GluLeuLysSerPheAspGlyAsp 153
Db 402 GACCGGGCCCGAGGCGACTGGCCACCGGCTCCAGAGCTGGAGGAGCGCGAGAGGCT 461
Qy 154 TyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu--- 172
Db 462 GCACGACGAGCGAGGAGGAATGAAGGTGATTGAGAACAGAGCGACGCAAAAGACGAGGAG 521
Qy 173 -----GluGlyLysLysArgLeuGluGluHisAsp 182
Db 522 AGATGCGAGATCCAGGAGATCGAGCTGAAGAGGCCCAAGCACATCCCGAGGAGCGCGAC 581
Qy 183 LysLysIleGluGluTyrLysLysGluIleThr-----MetAspLysAsn 197
Db 582 CGCAAGTACGAGGAGGTGGCTCGAAAACCTGGTATCCTCGAAGGAGACCTGGAGCGCTCC 641
Qy 198 GlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSer 217
Db 642 GAGGAACGCGCGAGGTGGCGGAGGCCAAATCTGGAGATCTT----- 683
Qy 218 TyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThr 237
Db 684 ----- 695
Qy 238 LysGlyLeuSerLysTyrLysLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSer 257
Db 696 AAAAAGCTCACCAACATCTTGAAGCCTCGGAAGCTCAG-----GCCGAGAAGTACTCC 749
Qy 258 GlnValAsn-----ProGluArgMetPheIleMetThrAsnLysAlaSerSerAsn 274
Db 750 CAAAAAGAGGACCAAGTATGAAGAGGAGATCCGAGTCTCTGACGAGCAAG----- 797
Qy 275 GluProSerLeuLysGluLeuGlyAspProValTyrLysLysLeuAsnAlaValLys 294
Db 798 -----CTGAGAGGCGGAGACTCGTGGGAGTTTCTGAGAGGTCGGTGCGCAAG 848
Qy 295 AsnGlnArg---ValAspIleLeuAspArgAspLeuTyrAlaArgSer---ArgGlyLeu 312
Db 849 CTGGAGAAGACCATCAGCATCTAGAAGAGCAGAGGTCTACGCTCAGAAAGCTGAAGGCGCAAG 908
Qy 313 IleSerSerGluGluMetAlaLysGluLeuValGluLeuSer 326
Db 909 GCTCTGAGCGAGGAGCTGGACCTGGCCCTGAATGACATGACC 950

RESULT 38
LOCUS CR685229 2252 bp mRNA linear HTC 19-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR685229
VERSION CR685229.1 GI:51183136
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
1 (bases 1 to 2252)
Genoscope.
REFERENCE Direct Submission
AUTHORS Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
TITLE Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
JOURNAL Submitted (10-AUG-2004) CP 5706 - 91057 EVRY cedex - FRANCE
(S-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
source
Location/Qualifiers
1..2252
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
```

```
ORIGIN
Alignment Scores:
Pred. No.: 0.157 Length: 2252
Score: 115.00 Matches: 70
Percent Similarity: 29.06% Conservatives: 58
Best Local Similarity: 21.88% Mismatches: 103
Query Match: 6.80% Indels: 92
DB: 3 Gaps: 16

US-10-724-972A-6352 (1-335) x CR685229 (1-2252)
Qy 34 SerSerLysGlyGluSerSerLysAspGlyValGluIleLysHisGlu-----Glu 49
Db 1068 TCTCAAAAGAGCGCCAAAGATGGATAAAGTCAAGATT---CAGGATATCGTGTGTCGCGC 1124
Qy 50 GlyThrThrLysValProLysHisProLysArgVal----- 61
Db 1125 GGTTCACCTCGCATCCCAAAATCCAGAAACTGTCTCCAGGACTACTTCAACGGGAGGTTT 1184
Qy 61 ----- 61
Db 1185 TGAACAAGAGCATCAACCTGATGAAGCTGTTGCTTATGTCGCCGCTGTCCAGGAGCCCA 1244
Qy 62 -----ValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysPro 79
Db 1245 TCTTGTCAAGTGACAAAGTCTGAGAATTGTGCAGGACCTGTCTCTCTGACGCTCACTCT 1304
Qy 80 ValGlyIleAlaAspAspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIle 99
Db 1305 CTGTCTCTTGGTATTGAGAGC---GCCGAGGCGTCATGACGCGTCTG-----ATC 1352
Qy 100 GlyLysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeu 119
Db 1353 AAGCCCAACACCACTCTTACCAAGCAGACCCAGACCTTCACCACTACTCTGACAC 1412
Qy 120 LysProAspLeuIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsn 139
Db 1413 CAGCCCGCGTCTCATCAGGTTTACGAGGGTGAGCGTGCATGACGAGGACCAACAT 1472
Qy 140 LysIle-----AlaProThrIle 145
Db 1473 CTGTCGGGCAAGTTCGAGTTGACGGGCATCCCCCTGCGCCCTCTGGGGGTTCCCCAGATC 1532
Qy 146 GluLeuLysSerPheAspGlyAspTyrAsn-----GluAsnIleAspAlaPheLysThr 163
Db 1533 GAGGTG---ACTTTTGACATCGACGCCCAACGGCATCATGAACGTGTCTGCC----- 1580
Qy 164 IleSerLysAlaLeuGlyLysGluGlu-----GlyLysLysArgLeu 178
Db 1581 GTGGACAAAAGCCTGGCAAGAAAACAAGATCACCATCACCACCAATGACAAAGGTGCTCTC 1640
Qy 179 GluGluHisAsp-----LysLysIleGluTyrLysLysGluIleThr 193
Db 1641 AGTAAGGAGGACATTGAGCGCATGTGCGAGGAAGTGAAGATCAAGGCCCAAGCAGAC 1700
Qy 194 MetAspLysAsnGlnLysValLeuProAlaValAlaAlaLysSerGlyLeuLeuAlaHis 213
Db 1701 GTCCAGCGTGACAAG-----GTGTCGGCCCAAGATGGCTTGGCTGAGTCTGATC 1745
Qy 214 ProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSer 233
Db 1746 GCTTTCAC-----ATGAAGTCAACCGGTGAA 1772
Qy 234 AspAspValThrLysGly---LeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsn 252
Db 1773 GACGAGAACTCGCTGGGAAGATCAGTGACGATGACGAGACGAGAGATTTGGACAAAGTGC 1832
Qy 253 ThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSer 272
Db 1833 AACGAGGTTCATCGCTGGCTGGCAAGATCAG-----ACTGCCGAGGAGGAC 1880
```

/tissue\_type="Byes"

```

QY 273 SerAsnGluProSerLeuLysGluLeuGluLys-----AspProValTrpLysLysLeu 290
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1881 GAGTACGAGCATCAGCAGAGGAGGTGGAGAGGTTGTCACACCCCATCATCACAAAGATG 1940

RESULT 39
BJ706325 707 bp mRNA linear EST 08-MAR-2004
LOCUS BJ706325 MF01FFA CDNA Oryzias latipes cDNA clone MF01FFA017e14 5',
DEFINITION mRNA sequence.
ACCESSION BJ706325
VERSION BJ706325.1 GI:45247204
KEYWORDS EST.
SOURCE Oryzias latipes (Japanese medaka)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
1 (bases 1 to 707)
REFERENCE Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
AUTHORS Medaka EST Project in Takeda's lab
TITLE Medaka EST Project in Takeda's lab
JOURNAL Unpublished (2001)
COMMENT Contact: Tadaou Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
Location/Qualifiers
Source 1..707
/organism="Oryzias latipes"
/mol_type="mRNA"
/strains="HG-rr"
/db_xref="taxon:8090"
/clone="MF01FFA017e14"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="fry stage 40"
/clone_lib="MF01FFA CDNA"

ORIGIN
Alignment Scores:
Pred. No.: 0.0352 Length: 707
Score: 114.50 Matches: 61
Percent Similarity: 42.23% Conservative: 45
Best Local Similarity: 24.30% Mismatches: 88
Query Match: 6.78% Indels: 57
DB: 4 Gaps: 13

US-10-724-972A-6352 (1-335) x BJ706325 (1-707)

QY 69 ValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLysLys 88
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
12 GTCCAAAGACTTCTGCTTCTGGATGTCACCTCCACTTCTCTGGGTATAGAGACT---GCT 68

QY 89 AsnArgIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArg 108
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
69 GGTGGGTGATGACAGTCTTA-----ATCAACGCAATACTACCAATTCCTACCAAA 119

QY 109 LysGlnProAsnLeuGluGluIleSerLysLysLeuLysProAspLeuIleAlaAspAsn 128
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
120 CAACCCAGACCTTTACTATACATCTGTATACAGCTGGGGTCTCATCCAGTTTAT 179

QY 129 AsnArgHisLysGlyIleTyrLysAspLeuAsnLysIle----- 141
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
180 GAAGGTGAGCGTCCATGACCAAGCAACCACTGCTGGGAAAGTTTGAACGACAGA 239

QY 142 -----AlaProThrIleGluLeuLysSerPheAspGlyAspTyr 154
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
240 ATTCTCTCTGCTTCCTCGTGGTGTCTCTCAGATTGAGGTG---ACATTGTATATTCACGA 296

QY 155 Aen-----GluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGlu 172

```

```

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
297 AATGAATCATGAATGTTCTGCT-----GTTGATAAGACACTGGCGCAAGAAAC 347

QY 173 Glu-----GlyLysLysArgLeuGluGluHisAsp----- 182
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
348 AAAATCAACATCACTAATAAGCAAAAGTCTGCTTTAGTAAAGAGGACATTTGAGCGCATGTC 407

QY 183 LysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuPro 202
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
408 CAAGAAGCTGAGAAGTACAAGGCAAGATGATGTTTCAGGTGTGACAA----- 455

QY 203 AlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe 222
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
456 ---GTGCTGCAAGAATGGCTGAGTCCATGCTTTCAC----- 494

QY 223 LeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGly---LeuSer 241
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
495 -----ATGAAGTCAACTGTGGAGATGAGAAAGCTTGTCTGGCAAGATCAGT 539

QY 242 LysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnPro 261
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
540 GATGAGGACAAACAAAGATCTTGGATAAGTGAATGAATGATCAGTGGCTGGACAAAG 599

QY 262 GluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeu 281
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
600 AACCAAG-----ACTGCTGAAGAGATGAGTACGAGCACCAACAGAGGAGCTT 647

QY 282 GluLys-----AspProValTrpLysLysLeu 290
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
648 GAGAAAGTATGTAATCCCATCATCACCAAACTG 680

RESULT 40
AL666132 711 bp mRNA linear EST 11-JAN-2002
LOCUS AL666132 directional larval cDNA library Ciona intestinalis cDNA
DEFINITION clone 002ZE10 5', mRNA sequence.
ACCESSION AL666132
VERSION AL666132.1 GI:18133039
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cionidae; Ciona.

REFERENCE 1 (bases 1 to 711)
AUTHORS Ciona intestinalis directional larval cDNA library
JOURNAL Unpublished (2002)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
IMPORTANT: this sequence may contain errors. The Ciona intestinalis
library from which the clone was isolated may be contaminated with
cDNAs from bacteria or other Eukarya.
Directional larval cDNA library originate from Dr.M.Branno,
Stazione A.Dohrn, Naples, Italy, and was prepared in
pBluescript2SK+.
FEATURES
Location/Qualifiers
Source 1..711
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="002ZE10"
/clone_lib="directional larval cDNA library"
/note="Vector: pBluescript2SK+"

ORIGIN
Alignment Scores:
Pred. No.: 0.0355 Length: 711
Score: 114.50 Matches: 67
Percent Similarity: 41.67% Conservative: 38
Best Local Similarity: 26.59% Mismatches: 89
Query Match: 6.78% Indels: 58

```



```
DB: 1 Gaps: 14
US-10-724-972A-6352 (1-335) x AL666132 (1-711)
QY 69 ValAspAlaValLeuValLeuAspValLysProValGlyIleAlaAspAsnLysLys 88
Db 19 GTTCAGGATCTTCTCTGTGGATGTGCACCTTTATCTCTGGTATTGAGACC---GCT 75
QY 89 AsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArg 108
Db 76 GGAGGAGTGATGACAGCTTTG-----ATCAAGAGAAACACGACTGTCCCAACCAAA 126
QY 109 LysGlnProAsnLeuGluIleSerLysLeuLysProAspLeuIleIleAlaAspAsn 128
Db 127 CAATCCCAATATTCACCACTTCTCGACAACCAACAGGTGTGTGATCCAAGTGTAC 186
QY 129 AsnArgHisLysGlyIleTyrLysAspLeuAsnLysIle----- 141
Db 187 GAGGGTCAACGTACGATGACCAAGGACAAACAACATACTTGGTAAGTTCGAGCTCTCCGGC 246
QY 142 -----AlaProThrIleGluLeuLysSerPheAspGlyAspTyr 154
Db 247 ATCCCCCTGCACCTCGTGGAGTTCACAGATTGAGGTG---ACGTTGATATCGATGCT 303
QY 155 AsnGlu-----AsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 172
Db 304 AACGGTATCTTGAATGTATCTGCT-----GTTGACAAGACACTGGAAAGGAGAAT 354
QY 173 Glu-----GlyLysLysArgLeuGluGluHisAsp----- 182
Db 355 AAGATCACANTACCAACGACAAGGTCTGCTTTAGCAAGAAGATATTGACGGATGGTG 414
QY 183 LysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuPro 202
Db 415 AACGAGCCCGAGAAATATAAGGAAGAA-----GATGAGAAGCAGAAG-----GAG 459
QY 203 AlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe 222
Db 460 AAGATTCAACGCAAGAACGGAACGACTCGAATCTTATGCGTTCAAC----- 501
QY 223 LeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGly---LeuSer 241
Db 502 -----CTTAATCAACGGTGGAGATGACAAAGTGAAGATAAGATTTC 546
QY 242 LysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnPro 261
Db 547 GACGAAGATAAATCAGCGATTCTTAAACAAGCTAAAGAGCTTTTAGACTGGTTGGAAAT 606
QY 262 Glu-ArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLe 281
Db 607 AACCAAGACTGCGGAAAGGACGAATATGAATTCCAACAAA-----AAAGAAGCT 654
QY 281 uGluLys-----AspProValTrpLysLysLeu 290
Db 655 TGAGAAAATTGCCAACCTTATTATGACAAAACCTT 688
```

Search completed: November 10, 2005, 00:11:33  
Job time : 3894 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 9, 2005, 19:57:58 ; Search time 935 Seconds  
(without alignments)  
2962.983 Million cell updates/sec

Title: US-10-724-972A-6352  
Perfect score: 1690  
Sequence: 1 GVESVRLKLTSLVIGLLFVL.....EEMAKELVELSKDKKDKK 335

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 9794790 seqs, 4134909567 residues

Total number of hits satisfying chosen parameters: 19589580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool\_h/US10724972/runat\_03112005\_141902\_29179/app\_query.fasta\_1.519  
-DB=Published Applications NA -QWTF=fastcap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10724972 @CGN 1.1.480 @runat\_03112005\_141902\_29179  
-NCPU=6 -ICPU=3 -NO MWAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOC=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELETE=7

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq\*
- 13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq2\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq\*
- 20: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq\*
- 21: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq\*
- 22: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq\*
- 23: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq\*
- 24: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq2\*
- 25: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq\*
- 26: /cgn2\_6/ptodata/1/pubpna/US11A\_NEW\_PUB.seq\*
- 27: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq\*
- 28: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1690	100.0	1008	24	US-10-724-972A-2580	Sequence 2580, Ap
2	1671	98.9	993	18	US-10-282-122A-34954	Sequence 34954, A
3	1242	73.5	1014	9	US-09-815-242-8156	Sequence 8156, Ap
4	1232	72.9	984	18	US-10-282-122A-7644	Sequence 7644, Ap
5	1181	69.9	927	9	US-09-815-242-4404	Sequence 4404, Ap
6	726	43.0	957	9	US-09-815-242-4194	Sequence 4194, Ap
7	726	43.0	960	9	US-09-815-242-8024	Sequence 8024, Ap
8	724.5	42.9	957	22	US-10-470-048B-301	Sequence 301, App
9	724.5	42.9	960	18	US-10-282-122A-7949	Sequence 7949, Ap
10	644	38.1	525	18	US-10-282-122A-35107	Sequence 35107, A
11	516.5	30.6	897	18	US-10-282-122A-30706	Sequence 30706, A
12	510.5	30.2	801	9	US-09-974-300-2066	Sequence 2066, Ap
13	475	28.1	668	8	US-08-781-986A-355	Sequence 355, App
14	475	28.1	668	19	US-10-329-624-355	Sequence 355, App
15	460	27.2	900	18	US-10-282-122A-19870	Sequence 19870, A
16	460	27.2	909	9	US-09-815-242-6372	Sequence 6372, Ap
17	460	27.2	909	18	US-10-282-122A-20621	Sequence 20621, A
18	460	27.2	10244	22	US-10-893-671-14	Sequence 14, Appl
19	454	26.9	891	18	US-10-282-122A-33236	Sequence 33236, A
20	405	24.0	990	22	US-10-470-048B-318	Sequence 318, App
21	405	24.0	999	17	US-10-278-946-15	Sequence 15, Appl
22	405	24.0	999	22	US-10-967-189-15	Sequence 15, Appl
23	405	24.0	3775	8	US-08-781-986A-238	Sequence 238, App
24	405	24.0	3775	19	US-10-329-624-238	Sequence 238, App
25	391	23.1	972	18	US-10-282-122A-9415	Sequence 9415, Ap
26	380.5	22.5	944	9	US-09-974-300-2061	Sequence 2061, Ap
27	375	22.2	2209	24	US-10-450-763-7183	Sequence 7183, Ap
28	375	22.2	4392	24	US-10-450-763-7183	Sequence 7183, Ap
29	350	20.7	2799	24	US-10-450-763-8842	Sequence 8842, Ap
30	334	19.8	2115	8	US-08-781-986A-13139	Sequence 13139, A
31	334	19.8	2115	19	US-10-329-624-604	Sequence 604, App
32	333	19.7	242	8	US-08-781-986A-2556	Sequence 2556, Ap
33	333	19.7	242	19	US-10-329-624-2556	Sequence 2556, Ap
34	331	19.6	945	9	US-09-738-626-375	Sequence 375, App
35	331	19.6	3309400	9	US-09-738-626-375	Sequence 375, App
36	329	19.5	648	9	US-09-974-300-2087	Sequence 2087, Ap
37	327	19.3	963	18	US-10-282-122A-17083	Sequence 17083, A
38	326	19.3	915	24	US-10-721-922A-351	Sequence 351, App
39	321	19.0	957	9	US-09-974-300-2044	Sequence 2107, Ap
40	317.5	18.8	321	9	US-09-815-242-2107	Sequence 2107, Ap
41	317.5	18.8	321	18	US-10-282-122A-4605	Sequence 4605, Ap
42	313	18.5	942	18	US-10-282-122A-10099	Sequence 10099, A
43	301	17.8	942	18	US-10-282-122A-9201	Sequence 9201, Ap
44	300.5	17.8	708	9	US-09-974-300-2068	Sequence 2068, Ap
45	296.5	17.5	1056	24	US-10-724-972A-3409	Sequence 3409, Ap

# ALIGNMENTS

## RESULT 1

US-10-724-972A-2580  
; Sequence 2580, Application US/10724972A  
; Publication No. US20040147734A1  
; GENERAL INFORMATION:  
; APPLICANT: Doucette-Stamm, Lynn  
; APPLICANT: Bush, David  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: PATH03-16  
; CURRENT APPLICATION NUMBER: US/10/724, 972A  
; CURRENT FILING DATE: 2003-12-01  
; PRIOR APPLICATION NUMBER: 09/450, 969  
; PRIOR FILING DATE: 1999-11-29

721 AGTAAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACCTTTATCTCAAGTGAAT 780  
261 ProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGlu 280  
781 CTGAGCGTATGTTTATATGACAAACAAAGCAAGTCTTAAGCAACCTTCACTTAAAGAA 840  
281 LeuGluLysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIle 300  
841 CTAGAAAAAGATCCTGTATGGAAGAAATTAACGCTGTGAAAAATCAACGTGTGTATTT 900  
301 LeuAspArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLys 320  
901 TTAGACCGTGACTTATGGGCAAGATCACGTGGTGGTAAATTTCTTCAGAGAATAATGGCAAAA 960  
321 GluLeuValGluLeuSerLysLysAspSerLysLysAspAsnLys 335  
961 GAACTGTTGAAATATCTAAGAAAGATAGTAAAAAAGATAATAAG 1005

RESULT 2  
US-10-724-972A-34954  
; Sequence 34954, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282.122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 34954  
; LENGTH: 993  
; TYPE: DNA  
; ORGANISM: Staphylococcus epidermidis  
US-10-724-972A-34954

Alignment Scores: 1.38e-135 Length: 1008  
Pred. No.: 1690.00 Matches: 335  
Score: 100.00% Conservativity: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 24 Gaps: 0  
DB: 24

US-10-724-972A-6352 (1-335) x US-10-724-972A-2580 (1-1008)

QY 1 GlyValGluSerValArgGlyLeuLysLysLeuSerValIleGlyLeuLeuPheValLeu 20  
DB 1 GGAGTGAATCAGTGAGAGGTTTAAAAATTTAAAGTGAATTTGGCTTATGTTTGTGTTTAA 60  
QY 21 IleAlaThrAlaAlaCysGlyAsnAsnSerSerSerSerSerSerLysLys 40  
DB 61 ATTGCAACTGCGCATGTGGAAATAATAGTTCAAGTAACCTCAAGTAAAGAGTCAATCAAAA 120  
QY 41 AspGlyValGluIleLysHisGluGluGlyThrThrLysValProLysHisProLysArg 60  
DB 121 GATGAGTTGAAATCAAGCACGAAGAGGTACTACGAAGTACCTAAACACCCCTAAAGCT 180  
QY 61 ValValValLeuGluTyrSerPheValAspAlaLeuAlaLeuAspValLysProVal 80  
DB 181 GTTGTGTTGCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTTAAACCTGT 240  
QY 81 GlyIleAlaAspAspAsnLysLysAsnArgIleLysProLysProLysLysLysGly 100  
DB 241 GGGATAGCGGATGATACAAAAAATCGTATTATTAACCATTAAGAGATAAATTTGGA 300  
QY 101 LysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluLysSerLysLysLys 120  
DB 301 AAATACACTTCTGTAGGAACAGTAAAGCAACCTTAAGTAAAGAAATCAGTAACTTAAA 360  
QY 121 ProAspLeuIleAlaAspAsnArgHisLysGlyIleTyrLysAspLeuAsnLys 140  
DB 361 CCAGATTTTAAATTTGCTGATTAATATAGACAAAGGTATTATTAAGACTTAAATAAA 420  
QY 141 IleAlaProThrIleGluLysSerPheAspGlyAspTyrAsnGluAsnIleAspAla 160  
DB 421 ATTGCTCTTACGATTTGAATGAAAGTTCGATGAGATTAATGAAATAATTTGATGCT 480  
QY 161 PheLysThrIleSerLysAlaLeuGlyLysGluGluGlyLysLysLysLysGluGlu 180  
DB 481 TTTAAACAATTTCAAAAGCTTTAGGTAAGAAGAAGAGGTAAAGGCTTTAGAGAA 540  
QY 181 HisAspLysLysLysLysLysLysLysGluIleThrMetAspLysAsnGlnLysVal 200  
DB 541 CACGATAAGAAAAATTTGAAGAATAATAAAAAAGAAATAATATGATGATTAATAATCAAAAGGTA 600  
QY 201 LeuProAlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGly 220  
DB 601 TTGCTCTGAGTAGCTGCTAATACAGGTTTGTGCTCATCCCAAGCAACTCTTATGTTGGT 660  
QY 221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAspValThrLysGlyLeu 240  
DB 661 CATTCTTAAGTCAACTAGGTTTTTAAAGAGCATTAAAGTATGATGTTACTTAAGGTTTAA 720  
QY 241 SerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsn 260

Alignment Scores: 5.83e-134 Length: 993  
Pred. No.: 1671.00 Matches: 331  
Score: 100.00% Conservativity: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00%

[illegible]



Db 541 ATCAATAAGTATAAAGATGAAATTAATTAATTTGATAGAAATCAAAAGTCTCCAGCAGTA 600  
Qy 205 AlalalysSerGlyLeuLeuAlaHisProSerAenSerTyrValGlyGlnPheLeuSer 224  
Db 601 GTTGCTAAAGCTGGTTTATTAGCACATCCAAACTATTTCATATGTTGGCAATTTTAAAC 660  
Qy 225 GlnLeuGlyPheLeuAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu 244  
Db 661 GAACATGAGATTAAATATGCAATTAAGTGACCATGTAACAAAGGTTTAAAGTAAATATTG 720  
Qy 245 LysGlyProTyrLeuGlnMetAenThrGluThrLeuSerGlnValAenProGluArgMet 264  
Db 721 AAAGGACCTTACTTACATTAAGACATCAACATTTAGCTGATTTAAATCCAGCGATG 780  
Qy 265 PheileMetThrAenLysAlaSerSerAenGluProSerLeuLysGluLeuLysAsp 284  
Db 781 ATCATATGACAGATCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
Qy 285 ProValTyrLysLeuAenAlaValLysAenGlnArgValAenPheLeuAspArgAsp 304  
Db 841 GCAACATGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
Qy 305 LeuTyrAlaArgSerArgGlyLeuLeuSerSerGluGluMetAlaLysGluLeuValGlu 324  
Db 901 GTTGGGCAAGATCTCTGCTGCTTAAATTTCTCTGGAAGATGCTAAAGAACTTTGTTGAA 960  
Qy 325 LeuSerLysLysAspSerLys 331  
Db 961 TTATCAAAAAGAACCAAAAG 981

## RESULT 5

US-09-815-242-4404  
; Sequence 4404, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE OF INVENTION: Prokaryotes  
; FILE REFERENCE: ELITRA.011a  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4404  
; LENGTH: 927  
; TYPE: DNA  
; ORGANISM: Staphylococcus aureus  
US-09-815-242-4404

Alignment Scores:

Pred. No.: 6.59e-92 Length: 927  
Score: 1181.00 Matches: 225

Percent Similarity: 85.11% Conservative: 38  
Best Local Similarity: 72.82% Mismatches: 46  
Query Match: 69.88% Indels: 0  
DB: 9 Gaps: 0

US-10-724-972a-6352 (1-335) x US-09-815-242-4404 (1-927)

Qy 16 LeuLeuPheValLeuLeuAlaThrAlaAlaCysGlyAenAenSerSerSerSerSerSer 35  
Db 1 TTATATGTTCCCTTATTTTATTTAGTTGCGAGCTTGTGGTAAATACGGATAATTCAGTAAACAAA 60  
Qy 36 LysGluSerSerLysAspGlyValGluLeuLysHisGluGluGlyThrThrLysValPro 55  
Db 61 GAATCATCACTAAAGATACTATTTCGTAAAGATGAAATCGTACAGTAAAGTAGTACCT 120  
Qy 56 LysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeu 75  
Db 121 AAAGATGCAAAACGATCTGTTGTTATAGAGTACTCATTTTCAGATGATGATGATGATGATG 180  
Qy 76 AspValLysProValGlyLeuAlaAspAspAsnLysLysAsnArgLysLysLysLysLys 95  
Db 181 GAGCTTAAACCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 240  
Qy 96 ArgAspLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 115  
Db 241 AGAGAAAAAATTCGGAATATATCTCTGTAGGTACAGTAAACAGCAAACTTAGAGGAA 300  
Qy 116 IleSerLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 135  
Db 301 ATTAGTAAATTAACACCGATTTAATATCGCTGATAGCAGTAGACATAGATTAAGGTATTAAT 360  
Qy 136 LysAspLeuAsnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 155  
Db 361 AAAGAAATTAACAAATTCGCAACATTTATCATTAAGAGTTTTCATGAGACTACAAA 420  
Qy 156 GluAsnIleAspAlaPheLysThrLysLysLysLysLysLysLysLysLysLysLysLys 175  
Db 421 CAACATTAATTCGTTCAACATTTGCTAAAGCTTTTAAATTAAGAAAGAAAGCGAA 480  
Qy 176 LysArgLeuGluGluHisAspLysLysLysLysLysLysLysLysLysLysLysLysLys 195  
Db 481 AAGCGTCTGCTGAACATGATATAATTAATCAAAAGATATAAGATGAATTAAGTTTGAAT 540  
Qy 196 LysAenGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSer 215  
Db 541 AGAAATCAAAAGTCTTCAGCAGTTGTTGCTAAAGCTGTTTATTAAGTATTAAGTTTGAAT 600  
Qy 216 AenSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAsp 235  
Db 601 TATTCATATGTTGGCAATTTTAAACGAACTTGGATTTAAATAATGCAATTAAGTATGAT 660  
Qy 236 ValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAenThrGluThr 255  
Db 661 GTAAACAAAGGTTTAAAGTAAATACTTTGAAGAGCCTTACTTACAATTAAGTATTAAGTAT 720  
Qy 256 LeuSerGlnValAenProGluArgMetPheLeuMetThrAenLysAlaSerSerSerSer 275  
Db 721 TTAGCTGACTTTAAATCCTGAACGATGATTTATGACAGATTAATGCTTAAAGAAAGATCT 780  
Qy 276 ProSerLeuLysGluLeuGluLysAspProValTyrLysLysLysLysLysLysLysLys 295  
Db 781 GCTGAATTCAGAAAGTTTACAAAGATCCAACTTGGAAAAAGTTGAAACGAGTTAAAAAT 840  
Qy 296 GlnArgValAspIleLeuAspArgAspLeuTyrAlaArgSerArgGlyLeuLeuSerSer 315  
Db 841 AATCGCGTGGATATTTGTTGACCGTGTGTTGGCAAGATCTCGTGGCTTAAATTTCTTCT 900  
Qy 316 GluGluMetAlaLysGluLeuValGlu 324  
Db 901 GAAGAAATGCTAAAGAACTTTGTTGAA 927

RESULT 6

US-09-815-242-4194

```

; Sequence 4194, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4194
; LENGTH: 957
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; US-09-815-242-4194

Alignment Scores:
Pred. No.:      8.32e-53      Length:      957
Score:          726.00      Matches:     153
Percent Similarity: 65.85%      Conservative: 61
Best Local Similarity: 47.08%      Mismatches:  99
Query Match:     42.96%      Indels:      12
DB:              9          Gaps:          5

US-10-724-972A-6352 (1-335) x US-09-815-242-4194 (1-957)

QY      6 ArgGlyLeuLysLeuSerValLleGlyLeuPheValLeuAlaThrAlaAla 25
Db      13 AGGAATATCGTTAAATAGTTGTGTTTATGCTTAATCTTCGTT---GTAGCAGTAGCGGGT 69
QY      26 CysGlyAsnAsnSerSerAsnSerSerLysGluSerSerLysAspGlyValGluLe 45
Db      70 TGTGTCATAAAGATGACT-----GAAGAAACTGAATGACGACAATA 114
QY      46 LysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeuGlu 65
Db      115 AAGATGATAGTAGGAACCTGAAATAATTAAGAAATACTTAAACGCTGTTGTTGTTATAGAA 174
QY      66 TyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyLeuAlaAspAsp 85
Db      175 TATAGTTTGTGCTGATTTATAGCAGCATTAGATATGAAACCTGTTGTTGTTGATGATGAT 234
QY      86 AsnLysLysAsnArgLleLleLysProLeuArgAspLysLysLysLysThrSerVal 105
Db      235 GGCAGCAGTAAATAATATAACAAGTCAAGTCAAGATAGAGTGGGCGCATATGATCGGTT 294
QY      106 GlyThrArgLysGlnProAsnLeuGluGluLysSerLysLeuLysProAspLeuLleLle 125
Db      295 GGAATCTAGCCGCAACCGAATATGGAAGTATGAAATAATTAACCGGATTTGATCAT 354
QY      126 AlaAspAsnAsnArgHisLysGlyLleTyrLysAspLeuAsnLysLleAlaProThrIle 145

```

```

Db      355 GCAGATGTTAGCAGACATAAGAAATCAAAATCAGAAATTAAGCAAAATTCGTCGACAATC 414
QY      146 GluLeuLysSerPheAspGlyAspTyrAsnGluAsnLysAspAlaPheLysThrIleSer 165
Db      415 ATGTTAGTTAGCGGTACGGGAGATTATAATGCAAAATATTGATGCAATTTAAACAGTCGCT 474
QY      166 LysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAspLysLysIle 185
Db      475 AAAGCAGTAGGCAAGAGAAAGCGAGAACCGTCTGGAAAGCATGATAAAATATTA 534
QY      186 GluGluTyrLysLysGluLleThrMetAspLysAsnGlnLysValLeuProAlaValAla 205
Db      535 GCGGAGATTAGAAAGAAATTTGAACAGAGTACGTTAAAACTGCAATTTGCAATTCGGTATC 594
QY      206 AlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGln 225
Db      595 TCAAGAGCAGGTATGTTTATTAATAATGAAGATACATTTATGGGACAAATTTCTTAATTA 654
QY      226 LeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLys 245
Db      655 ATGGGTATTCAACCTGAAAGTCACAAAAGACAAAACACTACGCATGTTGGTGAACGCAAGGGT 714
QY      246 GlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPhe 265
Db      715 GGTCTTATATTTTAAATAATGAAGAACTTGCCCAATATCAATCCAAAAGTTATGATT 774
QY      266 IleMetThrAsn---LysAlaSerSerAsnGluProSerLeuLysGluLeuLysAsp 284
Db      775 TTAGCCACTGACGGAACGACAAAATAGAACGAAATTC-----ATTGAT 822
QY      285 Pro---ValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspLleLeuAspArg 303
Db      823 CCTGCAGTTTGGAAATCATTAAAGCTGTGAAAGATAACAAAGTTTATGACGTTGACCGA 882
QY      304 AspLeuTrpAlaArgSerArgGlyLeuLysSerSerGluGluMetAlaLysGluLeuVal 323
Db      883 AATAAGTGGTTGAAATCAAGGGGTATTATCGCAAGTGAAGATATGGCAGAAAGATTAGAA 942
QY      324 GluLeuSerLysLys 328
Db      943 AAAATTGCAAAAAA 957

RESULT 7
; Sequence 8024, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22

```



13	ValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCysGlyAenAsnSerSerSer	32
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25	GTTGGTTTCATGCTAAATTTTAGTTGTAGCAGTACCGGGTTGGTCAAAAAGATACT---	81
26		
27		
28		
29		
30		
31		
32		
33	AsnSerSerLysGluSerSerLysAspGlyValGluIleLysHisGluGluGlyThrThr	52
34		
35		
36		
37		
38		
39		
40		
41		
42		
43		
44		
45		
46		
47		
48		
49		
50		
51		
52		
53	LysValProLysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeu	72
54		
55		
56		
57		
58		
59		
60		
61		
62		
63		
64		
65		
66		
67		
68		
69		
70		
71		
72		
73	ValAlaLeuAspValLysProValGlyIleAlaAspAsnLysLysAsnArgIleIle	92
74		
75		
76		
77		
78		
79		
80		
81		
82		
83		
84		
85		
86		
87		
88		
89		
90		
91		
92		
93	LysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsn	112
94		
95		
96		
97		
98		
99		
100		
101		
102		
103		
104		
105		
106		
107		
108		
109		
110		
111		
112		
113	LeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsnAsnArgHisLys	132
114		
115		
116		
117		
118		
119		
120		
121		
122		
123		
124		
125		
126		
127		
128		
129		
130	ATGGAAGTGATAAGTAAATTTAAACCGGATTTTCATCATTGCAGATGTTAGCAGACATAG	369
131		
132		
133	GlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGly	152
134		
135		
136		
137		
138		
139		
140		
141		
142		
143		
144		
145		
146		
147		
148		
149		
150		
151		
152		
153		
154		
155		
156		
157		
158		
159		
160		
161		
162		
163		
164		
165		
166		
167		

370 AAAATCAATCAGAAATTTGAGCAAAATTTGCTCCGACAATCATGTTAGTTCGGGTACGGGA 429  
153 AspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 172  
430 GATTATAATCAATATTTGAAGCATTTTAAACAGTCGCTTAAGCAGTAGGCAAGAGAAA 489  
173 GluGlyLysArgLeuGluGluHisAspLysIleGluGlyLysLysGluIle 192  
490 GAAGGCGAAGAGCGTCTGGAAAGCATGATAAATATTACCGGAGATTAGAAAGAAAT 549  
193 ThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAla 212  
550 GAACAGAGTACGTTAAATCTGCATTTTGCATTTTGCATTTTGCATTTTGCATTTTGCATTT 609  
213 HisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232  
610 AATAATGAAGATACATTTATGGGACAAATTCCTTAATTAATAATGGGTATTCACCTGAAGTC 669  
233 SerAspValThrLysGlyLeuSerLysTyrLysGlyProTyrLeuGlnMetAsn 252  
670 ACAAAAGACAAACTACGCATGTTGGTGAACGCGAAGGTCCTTATATATATATTTAAAT 729  
253 ThrGluThrLysSerGlnValAsnProGluArgMetPheIleMetThrAsn---LysAla 271  
730 AATGAAGAACTTGCCATATCAATCCAAAGATTATGATTTTAGCCACTGACGCAAAACG 789  
272 SerSerAsnGluProSerLeuLysGluLeuLysAspPro---ValTyrLysLeu 290  
790 GACAAAAATAGAACCAATTC-----ATTGATCTCGCAGTTTGGAAATCATTA 837  
291 AsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTyrAlaArgSerArg 310  
838 AAAGCTGTGAAGATACAAAGTTTATGACGTTGACCGAATAAGTGGTGAATCAAGG 897  
311 GlyLeuLysSerGluMetAlaLysGluLeuValGluLeuSerLysLys 328  
898 GGGATTATCGCAAGTGAAGTATGCGAAGATTTAGAAAAAATTTGCAGAAAAA 951

US-10-724-972A-7949  
; Sequence 7949, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR FILING DATE: 2003-02-20  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR FILING DATE: 2000-06-28  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR FILING DATE: 2000-03-47  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR FILING DATE: 2000-11-27

Alignment Scores: 1,12e-52 Length: 960  
Pred. No.: 724,50 Matches: 149  
Score: 66.67% Conservative: 63  
Best Local Similarity: 46.86% Mismatches: 95  
Query Match: 42.87% Indels: 11  
DB: 18 Gaps: 4

US-10-724-972A-6352 (1-335) x US-10-282-122A-7949 (1-960)

Qy 13 ValIleGlyLeuPheValLeuIleAlaThrAlaAlaCysGlyAsnAsnSerSerSer 32  
Db 25 GTTGTGTTTATGCTAAATCTTAGTTAGTACAGTAGCGGGTTGTGTCAAAAGATACT-- 81  
Qy 33 AsnSerSerLysGluSerSerLysAspGlyValGluIleLysHisGluGluThr 52  
Db 82 -----GAAGAGAAAACTGAAATGACGCAATAAAGATGAATTAGGAACTGAA 129  
Qy 53 LysValProLysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeu 72  
Db 130 AAAATTGAAGAAAAATCTTAACGTTGTTGTTAGTATATAGTATTTTGTGCTGATTATTA 189  
Qy 73 ValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLysLysAsnArgIle 92  
Db 190 GCAGCATTTAGATATGAAACCTGTTGGTATTGCAGATGATGCAGCACTTAAATAATAACA 249  
Qy 93 LysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsn 112  
Db 250 AAGTCAGTAAAGATAAGATTGGGGCATATGAATCGGTTGGATCTAGACCGCAACCGAAT 309  
Qy 113 LeuGluIleSerLysLeuLysProAspLeuIleAlaAspAsnAsnArgHisLys 132  
Db 310 ATGGAAGTGAATGAATTAACCGGATTTGATCAITTCAGATGTTAGCAGACATAAG 369  
Qy 133 GlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGly 152  
Db 370 AAAATCAATCAGAAATTTGAGCAAAATTTGCTCCGACAATCATGTTAGTTCGGGTACGGGA 429  
Qy 153 AspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 172  
Db 430 GATTATAATCAATATTTGAAGCATTTTAAACAGTCGCTTAAGCAGTAGGCAAGAGAAA 489  
Qy 173 GluGlyLysArgLeuGluGluHisAspLysIleGluGlyLysLysGluIle 192  
Db 490 GAAGGCGAAGAGCGTCTGGAAAGCATGATAAATATTACCGGAGATTAGAAAGAAAT 549  
Qy 193 ThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAla 212  
Db 550 GAACAGAGTACGTTAAATCTGCATTTTGCATTTTGCATTTTGCATTTTGCATTTTGCATTT 609  
Qy 213 HisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232  
Db 610 AATAATGAAGATACATTTATGGGACAAATTCCTTAATTAATAATGGGTATTCACCTGAAGTC 669  
Qy 233 SerAspValThrLysGlyLeuSerLysTyrLysGlyProTyrLeuGlnMetAsn 252  
Db 670 ACAAAAGACAAACTACGCATGTTGGTGAACGCGAAGGTCCTTATATATATATTTAAAT 729

[illegible]

RESULT 10

RESUB 10  
US-10-282-122A-35107  
; Sequence 35107, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.

```

/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/
/ FILE REFERENCE: ELITPA.034A
/
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/
/ CURRENT FILING DATE: 2003-02-20
/
/ PRIOR APPLICATION NUMBER: 60/191,078
/
/ PRIOR FILING DATE: 2000-03-21
/
/ PRIOR APPLICATION NUMBER: 60/206,848
/
/ PRIOR FILING DATE: 2000-05-23
/
/ PRIOR APPLICATION NUMBER: 60/207,727
/
/ PRIOR FILING DATE: 2000-05-26
/
/ PRIOR APPLICATION NUMBER: 60/230,335
/
/ PRIOR FILING DATE: 2000-09-06
/
/ PRIOR APPLICATION NUMBER: 60/230,347
/
/ PRIOR FILING DATE: 2000-09-09
/
/ PRIOR APPLICATION NUMBER: 60/242,578
/
/ PRIOR FILING DATE: 2000-10-23
/
/ PRIOR APPLICATION NUMBER: 60/253,625
/
/ PRIOR FILING DATE: 2000-11-27
/
/ PRIOR APPLICATION NUMBER: 60/257,931
/
/ PRIOR FILING DATE: 2000-12-22
/
/ PRIOR APPLICATION NUMBER: 60/267,636
/
/ PRIOR FILING DATE: 2001-02-09
/
/ PRIOR APPLICATION NUMBER: 60/269,308
/
/ PRIOR FILING DATE: 2001-02-16
/
/ Remaining Prior Application data removed - See file Wrapper or PALM.
/
/ NUMBER OF SEQ ID NOS: 78614
/
/ SOFTWARE: PatentIn version 3.1
/
/ SEQ ID NO 35107
/
/ LENGTH: 525
/
/ TYPE: DNA
/
/ ORGANISM: Staphylococcus haemolyticus
/
/ US-10-282-122A-35107

```

Alignment Scores:	
Pred. No.:	4, 4le-46
Score:	644.00
Percent Similarity:	84.00%
Best Local Similarity:	70.8%
Query Match:	38.11%
DB:	18
	4
	1
Length:	525
Matches:	124
Conservative:	23
Mismatches:	24
Indels:	4
Gaps:	1

```

US-10-724-972A-6352 (1-335) x US-10-282-122A-35107 (1-525)

Qy 5 ValArGlyLeuLyseIleuSerValIleGlyLeuLeuPheValLeuIleAlaThrAla 24
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 1 ATGAAGGCTTTAAAGTTTCAGGTATAGTCGCATATTATTATTTAGCAATAGTTTATTAGTTACT 60

Qy 25 AlaCysGlyAsnAsnSerSerSerAsnSer-----SerLyseIleuSerSerLyse 40
    : : : : : | | | | | : : : : : : : : : : : : : : : :
Db 61 GCTTGTGTAATGTTAGCAATAATGATCAGGTGATTTCAGGTAATAATATCGTCATCGAAA 120

Qy 41 AspGlyValGluIleLyseHisGluGlyThrThrLyseValProLyseHisProLyseArg 60
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 121 GATTCTATTAAAGTTTAAACACGCAATTAGGCACACTAAAGTTCTTAAAGATGCTAAACGT 180

Qy 61 ValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspVallyleProVal 80
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 181 GTCGTACGGTGGAAATTTTCATTTGTAGATGCTTTAGCGGCACATAAATGTTTAAACCTGTT 240

Qy 81 GlyIleAlaAspAsnLysLyseAsnArgIleIleLyseProLeuArgAspLysIleGly 100
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 241 GGGGTGCTGATGACAAACAAACCAATCGTATTATTAAACCATTAAAGAAANAATTTGGA 300

Qy 101 LysTyrThrSerValGlyThrArgLyseGlnProAsnLeuGluIleSerLyseLyse 120
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 301 GATTATAAACTGTGTGTGTCGTPAAACACACCCAACTTAGAAGAAATCAGTAAATTTAAA 360

Qy 121 ProAspLeuIleAlaAspAsnAsnArgHisLyseGlyIleTyrLyseAspLeuAsnLys 140
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 361 CCAGATTTAATATTGCAGATAGTAAACAGACACNAAGGTATTTCAAGAANAATTAAGTAAA 420

Qy 141 IleAlaProThrIleGluLeuLyseSerPheAspGlyAspTyrAsnGluAsnIleAspAla 160
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 421 ATTGCTCCACAATTCGAATTTAAAGAGTTTTCATGAGGAGATTATAATGATAATATAGATGCT 480

Qy 161 PheLysThrIleSerLyseAlaLeuGlyLyseGluGluGlyLys 175
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 481 TTTTAAACAATTCCAAAGCTTTTAAATAAAGATGATGTAGGTCCAG 525

```

## RESULT 11

```

RESULI 11
US-10-282-122A-30706
; Sequence 30706, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Cart, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625

```

281 LeuGluLysAspProValTrpLysLysLeuAenAlaValLysAenGlnArgValAspIle 300  
742 TGGGAAATGAAGCCTTATGGAAATATCCCTGGGTAATAATGGTCAAGTCATCTTA 801  
301 LeuAspArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLys 320  
802 GCTAATGATAATTTATGGCAAGACGCTGGTATTGATGCCGCTGAAGTAATGGCTAA 861  
321 GluLeuValGluLeuSerLysLysAspSerLys 331  
862 GAAGTCCAAGACTTTGTAAAGAAATCCGCCAA 894

RESULT 12  
US-09-974-300-2066  
; Sequence 2066, Application US/09974300  
; Patent No. US20020146721A1  
; GENERAL INFORMATION:  
; APPLICANT: Berka, Randy M.  
; APPLICANT: Clausen, Ib Groth  
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene  
; TITLE OF INVENTION: Expression  
; FILE REFERENCE: 10085.500-US  
; CURRENT APPLICATION NUMBER: US/09/974,300  
; PRIOR FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: 09/680,598  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/279,526  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 8481  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2066  
; LENGTH: 801  
; TYPE: DNA  
; ORGANISM: Bacillus licheniformis  
US-09-974-300-2066

Alignment Scores:  
Pred. No.: 2,17e-34 Length: 801  
Score: 510.50 Matches: 108  
Percent Similarity: 62.88% Conservative: 58  
Best Local Similarity: 40.91% Mismatches: 95  
Query Match: 30.21% Indels: 3  
Gaps: 3

US-10-724-972A-6352 (1-335) x US-09-974-300-2066 (1-801)

60 ArgValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysPro 79  
3 AAAGTTGGTCTTGAATTTGGGATTTTATGATGGCTCTTGTATGGCGGTATTAAAGCT 62  
80 ValGlyIleAlaAspAspAsnLysLysAenArgIleIleLysProLeuArgAspLysIle 99  
63 GTCCGAATTCGGACGACGCGGCAACCTAAGTTTATTACGAGAAGGTTCAGGGGAAAAATC 122  
100 GlyLysTyrThrSerValGlyThrArgLysGlnProAenLeuGluGluLysSerLysLeu 119  
123 AAAGGTATACCTTCACTTCAGTCGGTTCGGCGCCAGCAAGCTTTGAAAAAATTCCTCTTTA 182  
120 LysProAspLeuIleAlaAspAenAenArgHisLysGlyIleTyrLysAspLeuAen 139  
183 AAGCCGATTTAATTTATTCGCGACGACGCGGCGTCTATGATAGCTGTGCG 242  
140 LysIleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAenGluAenLys 159  
243 AAAATTCGCCGACCAATCCGCTCAAAATTTGAATCCGATTTATCAGGACACGCTTGTAT 302  
160 AlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGluLysLysArgLeuGlu 179  
303 GCATCTCTTACGATTTCGGAAGCGCTCGCAAGAAAGCAATGAGAAAAAATTTGGCT 362  
180 GluHisAspLysLysIleGluGluTyrLysLysGluIleThrMetAspLysAenGlnLys 199

PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 30706  
LENGTH: 897  
TYPE: DNA  
ORGANISM: Pasteurella multocida  
US-10-282-122A-30706

Alignment Scores:  
Pred. No.: 7.61e-35 Length: 897  
Score: 516.50 Matches: 111  
Percent Similarity: 60.48% Conservative: 65  
Best Local Similarity: 38.14% Mismatches: 98  
Query Match: 30.56% Indels: 17  
Gaps: 4

US-10-724-972A-6352 (1-335) x US-10-282-122A-30706 (1-897)

43 ValGluIleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValVal 62  
67 GTTACCGTCAAGATCAAAAGGGGAATTTACCTAGATAGCGTACCTAAACGTTGTGTT 126  
63 ValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIle 82  
127 GCCTTAGAATATCTTATGTGGATCGCACTTCACAAATTTGGTGCAGCCCTGTTGGTGTC 186  
83 AlaAspAenLysLysAenArgIleLysProLeuArgAspLysIleGlyLysTyr 102  
187 GCAGATGCAATGATAAAACCCGTTATCTGCAAAAGTACCGATTAAGTCGACCATGG 246  
103 ThrSerValGlyThrArgLysGlnProAenLeuGluIleSerLysLysLeuLysProAsp 122  
247 GAATCAGTGGGACGCGTCTCAACCGATTAGAGCGATTCTTCGCACTTAAACCCAGAT 306  
123 LeuIleAlaAspAenAenArgHisLysGlyIleTyrLysAspLeuAenLysIleAla 142  
307 TTGATCAITTCGGATGATAATCGCAATCTCTGCGCTATGAAGAACTCAAAAAATCGCG 366  
143 ProThrIleGluLeuLysSerPheAspGlyAspTyrAenGluAenLysIleAspAlaPheLys 162  
367 CCGACAGTCGTCTTAAATTCGCCCATGAGAACTATCAAGAAAACTTGAACCCGACAA 426  
163 ThrIleSerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluLysHisAsp 182  
427 AAATTCGGTGATTATTAGTAAATCAAAAGAAATGCAAGCGCTATTGCAAAACATAAG 486  
183 LysLysIleGluGluTyrLysLysGluIleThrMetAspLysAenGlnLysValLeuPro 202  
487 CAGGATATTTCGGACATCGCCAAA-----ACGTTACCGAAAGGGAAAAAGCGATTATC 540  
203 AlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAenSerTyrValGlyGlnPhe 222  
541 GGTGTTTCCAGTGAACCCAAATTTAATTTATATATAGCAATCCTATGCTGGTGCCTTA 600  
223 LeuSerGlnLeuGlyPheLys-----GluAlaLeuSerAspAspValThrLysGlyLeu 240  
601 GTGGAAGTGTAGTTTCAAAATGCCAAAGCCGCGTCCGATAAC----- 645  
241 SerLysTyrLeuLysGlyProTyrLeuGlnMetAenThrGluThrLeuSerGlnValAen 260  
646 -----CAACCTAATGCTTCGTTGGTTGGTTAGAACCAAGTGGCGGCGAGAAAG 690  
261 ProGluArgMetPheIleMetThrAenLysAlaSerSerAenGluProSerLeuLysGlu 280  
691 CCTGATCTGATGATCTTAATCCATTATTCGT-----GATGAGATATTGCAAGAAA 741

Db 363 GAACATAACAAAGCTGGATGATTAAGAAACAGAAATTCGGC--AGCCGGAACAGAGC 419  
 QY 200 ValLeuProAlaValAlaAlaLysSerGlyLeuAlaHisProSerAsnSerTyrVal 219  
 Db 420 ATTCTTCTGCTCGGAAACACAAATGACAAATCACCGTGGCGGATGAAACTTTTTCACG 479  
 QY 220 GlyLysPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGly 239  
 Db 480 TCTCAGCTTCTGACGAAATCGGTATACATACCGTGTGCGAGACAGC--GGCAAGGC 536  
 QY 240 LeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnVal 259  
 Db 537 GATGCCGAAACGGTGAATCGTCAATATTAAATGACGCTCGAACAGCTGCTTGAGAAA 596  
 QY 260 AsnProGluArgMetPheLeuMetThr---AsnLysAlaSerSerAsnGluProSerLeu 278  
 Db 597 GATCCCGACGCTTATTGTCTCATGACGGGAGAGAAAGATAAAGTCGATGAAGACGGGAAA 656  
 QY 279 LysGluLeuLysAspProValTyrLysLeuAsnAlaValLysAsnGlnArgVal 298  
 Db 657 AGACCATCGAAAGAGATCTCTTTGGAAAGAGCTCAGCGCAGTCAAAAAACGGCAAGGTT 716  
 QY 299 AspLeuAspArgAspLeuTyrAlaArgSerArgGlyLeuLysSerSerGluGluMet 318  
 Db 717 TATGAAGCCGACAGATTCGGCTCGCTCGGACGAGCATTCACGAGCAGATGAGCTG 776  
 QY 319 AlaLysGluLeu 322  
 Db 777 ATGATGAATC 788

RESULT 13

US-08-781-986A-355  
 ; Sequence 355, Application US/08781986A  
 ; Publication No. US20030054436A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Charles Kunsch  
 ; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
 ; NUMBER OF SEQUENCES: 5255  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Human Genome Sciences, Inc.  
 ; STREET: 9410 Key West Avenue  
 ; CITY: Rockville  
 ; STATE: Maryland  
 ; COUNTRY: USA  
 ; ZIP: 20850  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
 ; COMPUTER: HP Vectra 486/33  
 ; OPERATING SYSTEM: MSDOS version 6.2  
 ; SOFTWARE: ASCII Text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/781,986A  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Benson, Bob  
 ; REGISTRATION NUMBER: 30,446  
 ; REFERENCE/DOCKET NUMBER: PB248PP  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (301) 309-8504  
 ; TELEFAX: (301) 309-8512  
 ; INFORMATION FOR SEQ ID NO: 355:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 668 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; US-08-781-986A-355  
 Alignment Scores:

Pred. No.: 1.95e-31 Length: 668  
 Score: 475.00 Matches: 91  
 Percent Similarity: 86.89% Conservative: 15  
 Best Local Similarity: 74.59% Mismatches: 16  
 Query Match: 28.11% Indels: 0  
 DB: 8 Gaps: 0  
 US-10-724-972A-6352 (1-335) x US-08-781-986A-355 (1-668)  
 QY 210 LeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLys 229  
 Db 2 TTTATTTAGCACATCCAAACTATTTCATATGTTGGACAATTTTAAACGAACCTAGGATTTAAA 61  
 QY 230 GluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeu 249  
 Db 62 AATGCATTAAGTGACCATGTAAACAAAGGTTTAAGTAAATATTGAAAGGACCTTACTTA 121  
 QY 250 GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheLeuMetThrAsn 269  
 Db 122 CAATTAGACACTGAACATTTAGCTGATTAAATCCAGAGCGTATGATCATTTATGACAGAT 181  
 QY 270 LysAlaSerSerAsnGluProSerLeuLysGluLeuLysAspProValTyrLysLys 289  
 Db 182 CATGCTAAAAAGATTTCTGCTGAATTTCAAGAGTTTCAAGAGATGCAACATCGAAAAAG 241  
 QY 290 LeuAsnAlaValLysAsnGlnArgValAspLeuLeuAspArgAspLeuTyrAlaArgSer 309  
 Db 242 TTGAATGCAGTTAAAAAATAATCGCGTGGATATTGTTGACCGTGATGTTGGGCAAGATCT 301  
 QY 310 ArgGlyLeuLysSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAsp 329  
 Db 302 CGTGCTTAATTTCTCTCGAAGAAATGGCTAAAGAACTTGTGTAATTTATCAAAAAAGAA 361  
 QY 330 SerLys 331  
 Db 362 CAAAAG 367  
 RESULT 14  
 US-10-329-624-355  
 ; Sequence 355, Application US/10329624  
 ; Publication No. US20040043037A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Charles Kunsch  
 ; Gil H. Choi  
 ; Patrick S. Dillon  
 ; Craig A. Rosen  
 ; Steven C. Barash  
 ; Michael R. Fannouh  
 ; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
 ; NUMBER OF SEQUENCES: 5256  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Human Genome Sciences, Inc.  
 ; STREET: 9410 Key West Avenue  
 ; CITY: Rockville  
 ; STATE: Maryland  
 ; COUNTRY: USA  
 ; ZIP: 20850  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
 ; COMPUTER: HP Vectra 486/33  
 ; OPERATING SYSTEM: MSDOS version 6.2  
 ; SOFTWARE: ASCII Text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/10/329,624  
 ; FILING DATE: 27-Dec-2002  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/956,171  
 ; FILING DATE: October 20, 1997  
 ; APPLICATION NUMBER: 60/009,861  
 ; FILING DATE: January 5, 1996  
 ; APPLICATION NUMBER: 08/781,986  
 ; FILING DATE: January 3, 1997  
 ; ATTORNEY/AGENT INFORMATION:

```

; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248PID1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 355:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 668 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 355:
US-10-329-624-355

Alignment Scores:
Pred. No.: 1,95e-31 Length: 668
Score: 475.00 Matches: 91
Percent Similarity: 86.89% Conservative: 15
Best Local Similarity: 74.59% Mismatches: 16
Query Match: 28.11% Indels: 0
DB: 19 Gaps: 0

US-10-724-972A-6352 (1-335) x US-10-329-624-355 (1-668)
QY 210 LeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLys 229
Db 2 TTATTAGCACATCCAAACTATTTCATATGTTGGCAATTTTAAACGAACACTAGGATTTAAA 61
QY 230 GluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeu 249
Db 62 AATGCATTAAGTCAGCATGTAACAAAGGTTTAACTAAATATTTGAAAGGACCTTACTTA 121
QY 250 GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheLeuMetThrAsn 269
Db 122 CAATTAGACACTGAACATTTAGCTGATTTAAATCCAGACGCTATGATCATTTATGACAGAT 181
QY 270 LysAlaSerSerAsnGluProSerLeuLysGluLeuGluLysAspProValTyrLysLys 289
Db 182 CATGCTAAAAAGATTTCTGCTGAATTCAGAAGATTACAAGAGATGCAACATGGAAAG 241
QY 290 LeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTyrAlaArgSer 309
Db 242 TTGAATGCAGTTAAAAATAATCGCGTGGATATTGTTGACCGTGATGTTGGCGAAGATCT 301
QY 310 ArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAsp 329
Db 302 CTGGCTTAATTTCTTCTGAAGAAATGCTTAAGAACTTTGTTGAATTATCAAAAAAGAA 361
QY 330 SerLys 331
Db 362 CAAAG 367

RESULT 15
US-10-282-122A-19870
; Sequence 19870, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
;
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19870
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Enterobacter cloacae
US-10-282-122A-19870

Alignment Scores:
Pred. No.: 5.45e-30 Length: 900
Score: 460.00 Matches: 101
Percent Similarity: 57.09% Conservative: 60
Best Local Similarity: 35.82% Mismatches: 105
Query Match: 27.22% Indels: 16
DB: 18 Gaps: 4

US-10-724-972A-6352 (1-335) x US-10-282-122A-19870 (1-900)
QY 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValLeu 64
Db 70 GTTCAAGACGACACGCGACGCTTTACACTCGATAAACCCACACGATGTTGGTGTG 129
QY 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyLeuAlaAsp 84
Db 130 GAATCTCTGTTGCGCGATGCGCTGCGCGCGTGCAGCGTCAGCGCGATCGTATTGCCGAC 189
QY 85 AspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSer 104
Db 190 GATAACGATGCAAAACGATCTCTGCCGAGTGGCTGCGACCTGAAACCCGTGGCAGTCC 249
QY 105 ValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIle 124
Db 250 GTCGGAACGCGCGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 309
QY 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144
Db 310 ATTGCCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 369
QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluLysAsnIleAspAlaPheLysThr 164
Db 370 CTGCTCTTAAGTCCCGCAACGAAACCTACGCTGAAATTTTGCACCTCAGCGCGTATCATC 429
QY 165 SerLysAlaLeuGlyLysGluGlyLysLysLysLysLysLysLysLysLysLysLys 184
Db 430 GCGCAAGTGGTAGGTAATAAAACGAGAGATGCGCGCGCGCGCGCGCGCGCGCGCG 489
QY 185 IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204
Db 490 ATGGCGCAGTGGTCCAGCCAG-----CTTCCCACAGGACACGCGCGCTTTGGCACA 543
QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224
```

Db 544 TCACGCGAAGACAGTATTTCACCTGCATACCCAGAGAGACCTGGAGCCGGAGCGTCTGGCT 603  
QY 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu 244  
Db 604 TCTCTGGGGCTG-----AACGTTCCCGCTGCGATG 633  
QY 245 LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263  
Db 634 GCGGGCCGCTCCATGCGCTCCATCGGCTGGAGCAACTGCTGGCGGTCAATCCTCGCTGG 693  
QY 264 MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLys 283  
Db 694 CTGCTGGTGGCCCACTATCG-----GAAGAGAGCATTTGTTAAACGCTGGCAACAA 744  
QY 284 AspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg 303  
Db 745 GATCCGCTCTGGCAGATATTAAACCGCGCAGAGCAGCAGGTTCCTTCAGTCGACAGT 804  
QY 304 AspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuVal 323  
Db 805 AACGCTGGCGCGGATGCGCGGTATTTTTCGCGAGAGCGTATTCCGCTGACAGCGTA 864  
QY 324 GluLeu 325  
Db 865 AAAATC 870

RESULT 16

US-09-815-242-6372, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIORITY APPLICATION NUMBER: 60/191,078  
; PRIORITY FILING DATE: 2000-03-21  
; PRIORITY APPLICATION NUMBER: 60/206,848  
; PRIORITY FILING DATE: 2000-05-23  
; PRIORITY APPLICATION NUMBER: 60/207,727  
; PRIORITY FILING DATE: 2000-05-26  
; PRIORITY APPLICATION NUMBER: 60/242,578  
; PRIORITY FILING DATE: 2000-10-23  
; PRIORITY APPLICATION NUMBER: 60/253,625  
; PRIORITY FILING DATE: 2000-11-27  
; PRIORITY APPLICATION NUMBER: 60/257,931  
; PRIORITY FILING DATE: 2000-12-22  
; PRIORITY APPLICATION NUMBER: 60/269,308  
; PRIORITY FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6372  
; LENGTH: 909  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(909)  
US-09-815-242-6372

Alignment Scores: 5.52e-30 Length: 909  
Pred. No.: 460.00 Matches: 100  
Score:

Percent Similarity: 57.45% Conservative: 62  
Best Local Similarity: 35.46% Mismatches: 104  
Query Match: 27.22% Indels: 16  
DB: 9 Gaps: 4  
US-10-724-972A-6352 (1-335) x US-09-815-242-6372 (1-909)  
QY 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64  
Db 76 GTTCAGAGCAACACGCGCTTTACACTCGAAAAACGCCAACAGGATTTGGTGTCTG 135  
QY 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84  
Db 136 GAACTCTCGTTCCCGATGCGCTGGCGCGCTGAGCTCATCCCGATTCGGTATTCGCAC 195  
QY 85 AspAsnLysLysAsnArgIleLysProLeuArgAspLysIleGlyLysTyrThrSer 104  
Db 196 GATAACGATGCAAAACGATCTCTGCCGAAGTCGTCGCCACCTGAAACCGTCAGTC 255  
QY 105 ValGlyThrArgLysGlnProAsnLeuGluLysSerLysLysLysProAspLeuIle 124  
Db 256 GTCCGAGACGCGCGCGAGCGCGAGCTTGGAGACCTTGGCGCTCTGAAACGACACCTGATC 315  
QY 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144  
Db 316 ATTGCCAGACGAGTCGCTATCGGGGGTTTACATCGCTTGCAGCAATTCGCGCGGTA 375  
QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThr 164  
Db 376 CTGCTGCTTAAGTCCCGCAACGAAACCTACGCTGAAATTTGCAATCTCGCGCTATCATC 435  
QY 165 SerLysAlaLeuGlyLysGluGlyLysLysArgLeuLeuGluLysLysLys 184  
Db 436 GCGCAATGCTGGTAAAGCAGAGATGACGAGCAGCTCTGGAACAAACATAAGAGAGG 495  
QY 185 IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204  
Db 496 ATGGCGCAGTGGCGCCAGCAG-----CTTCCCAAGGGACACGCGTGGCCTTTGGCACA 549  
QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224  
Db 550 TCACGCGAAGACCAATTCAACCTGCATCTCAGAGAGACCTGGACCGCAGCGTGTGGCC 609  
QY 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu 244  
Db 610 TCTCTGGGGCTG-----AACGTTCCCGCTGCGATG 639  
QY 245 LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263  
Db 640 GCGGTGCGTCCATCGCTCCATCGGCTGGAGCACTGCTGGCGGTCAATCCTCGCTGG 699  
QY 264 MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLysGluLeuGluLys 283  
Db 700 CTGCTGGTTGCCCACTATCGC-----GAAGAGAGCATTTGTTAAACGCTGGCAACAA 750  
QY 284 AspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg 303  
Db 751 GATCCGCTCTGGCAGATGTTAAACCGCGCGCAGAGCAGCAGGTTCGCTTCGCTGACAGT 810  
QY 304 AspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuVal 323  
Db 811 AACACCTGGCGCGGATGCGCGGTATTTTTCGCGAGAGCGTATTTCGCGCTGACACGTA 870  
QY 324 GluLeu 325  
Db 871 AAAATC 876

RESULT 17

US-10-282-122A-20621  
; Sequence 20621, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 20621  
LENGTH: 909  
TYPE: DNA  
ORGANISM: Escherichia coli  
US-10-282-122A-20621  
Alignment Scores:  
Pred. No.: 5,52e-30 Length: 909  
Score: 460.00 Matches: 100  
Percent Similarity: 57.45% Conservative: 62  
Best Local Similarity: 35.46% Mismatches: 104  
Query Match: 27.22% Indels: 16  
DB: 18 Gaps: 4  
US-10-724-972A-6352 (1-335) x US-10-282-122A-20621 (1-909)  
QY 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64  
Db 76 GTTCAGGACGACACGGCAGTTTACACTCGAAAAAAGCCACACGGATGTGTGTG 135  
QY 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84  
Db 136 GAATCTCTGTCGCGAGCGTGGCGCGCGGAGCGTATCCGATCGGTATTCGGAC 195  
QY 85 AspAsnLysLysAsnArgIleIleLysProLeuArgAspLysLysLysThrSer 104  
Db 196 GATAACGATGCAAAAGCCATCTGCCCGAAGCGTGGCGACACCGTGGCGAGTCC 255  
QY 105 ValGlyThrArgLysGlnProLeuGluGluLeuSerLysLysLysProAspLeuIle 124  
Db 256 GTCCGAACCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 315  
QY 125 IleAlaAspAsnArgHisLysGlyIleTyrLysAspLeuAsnLysLysLysLeuProThr 144  
Db 316 ATTCGCCGACGACGAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 375

QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164  
Db 376 CTGCTGCTTAAGTCCCGCAACGAACCTACGCTGAAAAATTTGCAATCTGGCGGTATCATC 435  
QY 165 SerLysAlaLeuGluGlyLysGluGluGlyLysArgLeuGluGluHisAspLysLys 184  
Db 436 GCGCAATGTTGGTAAAGCGAGATGCGACGCGTCTGGAACAACATAAAGAGAGG 495  
QY 185 IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204  
Db 496 ATGGCGAGTGGCGCCAGCCAG-----CTTCCCAAGGGACACGCGTGGCGCTTTGGCACA 549  
QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224  
Db 550 TCACGCAACAGCAATTTCAACCTGCATCTACGAGACCTGGACCGCGCTGCTGGCC 609  
QY 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysLysLeu 244  
Db 610 TCTCTGGGCGTG-----AACGTTCCCGTGGCGATG 639  
QY 245 LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263  
Db 640 GCGGTCGTCCTCATCGCGCTCATCGGCTGAGCAACTGTGGCGGTCAATCTCTGCTGG 699  
QY 264 MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLys 283  
Db 700 CTGCTGTTGCCACTATCG-----GAAGAGAGCATTTGTTAAACGCTGGCAACA 750  
QY 284 AspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg 303  
Db 751 GATCGCTCTGGCAGATGTTAACCGCGCGCGAGCAGCAGGTGCTTCGTCGACAGT 810  
QY 304 AspLeuTrpAlaArgSerArgGlyLeuIleSerGluGluMetAlaLysGluLeuVal 323  
Db 811 AACACCTGGCGCGAGTCTGCGGTATTTTCTGCGAGCGTATTCGCGTGACACGTA 870  
QY 324 GluLeu 325  
Db 871 AAAATC 876  
RESULT 18  
US-10-893-671-14/c  
; Sequence 14, Application US/10893671  
; Publication No. US20050064527A1  
; GENERAL INFORMATION:  
; APPLICANT: Levy, Stuart, et. al.  
; TITLE OF INVENTION: NMR COMPOSITIONS AND THEIR METHODS OF USE  
; FILE REFERENCE: PKZ-043  
; CURRENT APPLICATION NUMBER: US/10/893,671  
; CURRENT FILING DATE: 2004-07-15  
; PRIOR APPLICATION NUMBER: US/09/801,563  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: 60/188,362  
; PRIOR FILING DATE: 2000-03-10  
; NUMBER OF SEQ ID NOS: 98  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 10244  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
US-10-893-671-14  
Alignment Scores:  
Pred. No.: 1.07e-28 Length: 10244  
Score: 460.00 Matches: 100  
Percent Similarity: 57.45% Conservative: 62  
Best Local Similarity: 35.46% Mismatches: 104  
Query Match: 27.22% Indels: 16  
DB: 22 Gaps: 4  
US-10-724-972A-6352 (1-335) x US-10-893-671-14 (1-10244)



```
QY 45 IleYHisGluGluGlyThrLysValProLysHisProLysArgValValValLeu 64
Db 7735 GTTCAGGACGACGACGACGTTTACACTCGAAACACGACACGATGTGGTCTG 7676
QY 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84
Db 7675 GAATCTCTGTTGGCGGATGGCTGGCGCGTGGACGTCATCCGATCGGTATGCCGAC 7616
QY 85 AspAsnLysLysAsnArgGlyIleLysPheProLeuArgAspLysLysGlyLysTyrThrSer 104
Db 7615 GATAACGATGCAAAACCATCTCTCCGGAAGTGGTGGCGACCTCGCAACCGTGGCGATCC 7556
QY 105 ValGlyThrArgLysGlnProAsnLeuGluGluLysSerLysLysLysProAspLeuLeu 124
Db 7555 GTCCGACGCGCGGACGCGGAGCTGGAAGCATTTGCCCTCTGAAACGACGACCTGATC 7496
QY 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysLysIleAlaProThr 144
Db 7495 ATTGCGGACAGCATGCGCATGGCGGGTTTACATCGCTTGCAGCAATCGCGCGGTA 7436
QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164
Db 7435 CTCTCTCTTAAGTCCCGCAGCAAACTACGCTGAAATTTGCAATCTGCGGCTATCATC 7376
QY 165 SerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAspLysLys 184
Db 7375 GCGCAATGTGGTGAAGACGAGATGCGGCGACGTCGTGAAACAATGAGAGAGG 7316
QY 185 IleGluGluTyrLysLysGlyIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204
Db 7315 ATGGCGCAGTGGGCCAGCCAG-----CTTCCAAAGGACGACGCGTGGCCCTTGGCACA 7262
QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224
Db 7261 TCACGCGAAGCAGCAATTAACCTGCATCTCAGGAGACCTGGACCGGACGCTGTGGCC 7202
QY 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu 244
Db 7201 TCTCTGGGCTG-----AACGTTCCCGCTGCGATG 7172
QY 245 LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263
Db 7171 GCGGGTGGCTCCATCGCGTCCATCGCGCTGGAGCACTGCTGGCGGTCAATCTGCTGG 7112
QY 264 MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGlyLys 283
Db 7111 CTGCTGGTGGCCACTATCGC-----GAAGAGAGCATTTGTTAAACGCTGGCAACA 7061
QY 284 AspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg 303
Db 7060 GATCCGCTCTGGCAGATGTTAACCCGCGCGCAGAGACGAGTGTGCTTGGTTCGACAGT 7001
QY 304 AspLeuTrpAlaArgSerArgGlyLeuLysSerSerGluGluMetAlaLysGluLeuVal 323
Db 7000 AACACCTGGCGCGATGCGCGGTATTTTGTGCGAGAGCGTATTTGCCGCTGACACGTA 6941
QY 324 GluLeu 325
Db 6940 AAAATC 6935
```

## RESULT 19

US-10-282-122A-33236

; Sequence 33236, Application US/10282122A

; Publication No. US20040029129A1

## GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

```
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 33236

; LENGTH: 891

; TYPE: DNA

; ORGANISM: Pseudomonas syringae

US-10-282-122A-33236

## Alignment Scores:

Pred. No.:	1,76e-29	Length:	891
Score:	454.00	Matches:	110
Percent Similarity:	52.80%	Conservative:	60
Best Local Similarity:	34.16%	Mismatches:	104
Query Match:	26.86%	Indels:	48
DB:	18	Gaps:	7

US-10-724-972A-6352 (1-335) x US-10-282-122A-33236 (1-891)

```
QY 17 LeuPheValLeuIleAlaThrAlaAlaCysGlyAsnAsnSerSerSerAsnSerLys 36
Db 13 CTGCTCAGCTGTGGCGAGCGCGCT----- 39
QY 37 GluSerSerLysAspGlyValGluIleLysHisGluGluGlyThrLysVal----- 54
Db 40 -----CAGGCTGCGCCCATCGACATCGATGACGCGCACGACCAAGTGCATCTG 87
QY 55 ProLysHisProLysArgValValValValLysSerPheValAspAlaLeuValala 74
Db 88 CCAGACACCCCAAGCGCGTGTGTAATGCTGAAATTTCTGCGACGGGCTTGTCTCG 147
QY 75 LeuAspValLysProValGlyIleAlaAspAspAsnLysLysAsnArgIleLysPro 94
Db 148 GTTGGCGTGACACCGGTGCGCGCGCGAGATGATGGCGACGCGCAGTGTGTACTGCCCAAG 207
QY 95 LeuArgAspLysLysLysTyrThrSerValGlyThrArgLysGlnProAsnLeuGlu 114
Db 208 GTGCGCAAGCGGTGGGTGTAATGTCAGTGGTGGGCTGCGCTCGCAACCAATATCGAA 267
QY 115 GluLysSerLysLysProAspLeuIleAlaAspAsnAsnArgHisLysGlyIle 134
Db 268 GTGATCGCAGCGCTCAAGCGGACCTGATCATCGCTCGGTCGTCATCAGGCTCTT 327
QY 135 TyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGlyAspTyr 154
```

Db	328	TATAACGACCTGCCAGCCTGGCTCGACGCTGATGTCGGTCACGCGCGAGGACTAT	387
Qy	155	AsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGly	174
Db	388	CAGGCGAGCTGAAATCTGCGCGCTGATCGGATCGCGTGGCGAAAGCCGCGAGATG	447
Qy	175	LysLysArgLeuGluGluHisAspLysLysIleGluGluTyrlsLysGluIleThrMet	194
Db	448	CAGGCGGAATCGCAGAAACCGTCAGACCTGAAACACCTTCCCGAGCAGATCTCGCC	507
Qy	195	AspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisPro	214
Db	508	GACAGCAAC-----GTGCTGTTCCGGTTCGCGTGAACACAGCTTCTCGGTCACGCG	561
Qy	215	SerAsnSerTyrlsValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAsp	234
Db	562	CCGCACTCTACGCGAGCGCGTGTCAAGCCATCGT-----	600
Qy	235	AspValThrLysGlyLeuSerLysTyrlsLeuLysGlyProTyrlsLeuGlnMetAsnThr---	253
Db	601	-----TTGCAAGTACCGGAAGTGCACCAACACGCGCG	633
Qy	254	-----GluThrLeuSerGlnValAsnProGluArgMetPheIle	266
Db	634	CCCACCGAGTTCGTCAGCTGGACCACTGCTCCCTCGGACCCCAACTGGTTCGTGTC	693
Qy	267	MetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeu---GluLysAspPro	285
Db	694	-----GGCCATTATCGTCGCGCGAGCATTTGTGCACACCTGCGAGCAAGCAGCGG	741
Qy	286	ValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeu	305
Db	742	CTGTGGCAGGTGCTCGGCGCTGTACGTAAATAAACAGGTGGCGGAGTCGCGGTGATAGC	801
Qy	306	TrpAlaArgSerArgGlyLeuIleSerGluGluMetAlaLysGluLeuValGluLeu	325
Db	802	TGGCGCGCAATCGCGGAATCATGCGTCCGAGCGATCGCTGACGATGCGCTGCGCGTA	861
Qy	326	SerLys 327	
Db	862	CTCAAG 867	
RESULT 20			
US-10-470-048B-318			
; Sequence 318, Application US/10470048B			
; Publication No. US2005003744A1			
; GENERAL INFORMATION:			
; APPLICANT: MEINKE ET AL.			
; TITLE OF INVENTION: A METHOD FOR IDENTIFICATION, ISOLATION AND PRODUCTION OF			
; FILE OF INVENTION: ANTIGENS TO A SPECIFIC PATHOGEN			
; FILE REFERENCE: SONN:035US			
; CURRENT APPLICATION NUMBER: US/10/470,048B			
; CURRENT FILING DATE: 2003-07-25			
; NUMBER OF SEQ ID NOS: 603			
; SOFTWARE: Patentin version 3.1			
; SEQ ID NO 318			
; LENGTH: 990			
; TYPE: DNA			
; ORGANISM: Staphylococcus aureus			
US-10-470-048B-318			
Alignment Scores:			
Pred. No.: 3,25e-25 Length: 990			
Score: 405.00 Matches: 113			
Percent Similarity: 51.74% Conservative: 65			
Best Local Similarity: 32.85% Mismatches: 132			
Query Match: 23.96% Indels: 34			
DB: 22 Gaps: 12			
US-10-724-972A-6352 (1-335) x US-10-470-048B-318 (1-990)			
Qy	8	LeuLysIleLeuSerValIleGlyLeuPheValLeuIleAlaThrAlaLysGly	27

RESULT 21  
US-10-278-946-15  
; Sequence 15, Application US/10278946

Db	541	TTCCAAAAA-----GATGCAAAAGCAAGATATAAAGATGCATGGCCATTGAA	588
QY	208	SerGlyLeuLeuAlaHisProSerAen-----SerTyrValGly	220
Db	589	GCTTCAGTTGTAACTTCGCTGCTGATCATACAAGNAITTTATGCTGGTGGATATGCTGGT	648
QY	221	GlnPheLeuSerGlnLeuGlyPheIysGluAlaLeuSerAspValThrLysGlyLeu	240
Db	649	GAATCTTAAATGATTTAGGATTTCAAACGT-----AATAAAGACTTA	690
QY	241	SerLysTyrLeuLysGly-----ProTyrLeuGlnMetAsnThr---GluThrLeuSer	257
Db	691	CAAAAACAAGTTGATAATGGTTAAAGATATATCCAACTTACATCTAAAGAAAGCATTC	750
QY	258	GlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSer	277
Db	751	TTAATGAACGCTGATCATATTTTCTAGTAAATCAGATCCAAATGCGAAAGAGTGTGCA	810
QY	278	Leu-----LysGluLeuGluLysAspProValTrpLysLysLeuAsnAla	292
Db	811	TTAGTTAAAAAGACTGAAGCGCAATGGACTTCAAGTAAAGAGTGGAAAAAATTTAGACGCA	870
QY	293	ValLysAsnGlnArgVal---AspIleLeuAspArgAspLeuTrpAlaArgSerArgGly	311
Db	871	GTTAAAAACAACCAAGTATCTCGATGATTTAGATGAATCACCITGGAACATTGCTGCGCGA	930
QY	312	LeuIleSerSerGluGluMetAlaLysGluValGluLeuSerLysLysAspSerLys	331
Db	931	TATAAATCTTATAAAACTTATTCACGATTTATATGAA-----AAGTTAAATATTGAA	984
QY	332	LysAspAsnLys	335
Db	985	AAACAATCAAAA	996

```

RESULT 22
US-10-967-189-15
; Sequence 15, Application US/10967189
; Publication No. US20050053995A1
; GENERAL INFORMATION:
; APPLICANT: Simpson et al.
; TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PR461USD2
; CURRENT APPLICATION NUMBER: US/10/967,189
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: US 10/278,946
; PRIOR FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: US 09/830,217
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: PCT/US99/06199
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: US 60/078,682
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 60/080,296
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: US 60/084,674
; PRIOR FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 999
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-967-189-15

```

Alignment Scores:	
Pred. No.:	3,28e-25
Score:	405.00
Length:	999
Percent Similarity:	51.74%
Matches:	113
Percent Local Similarity:	32.85%
Conservative:	65
Query Match:	23.96%
Mismatches:	132
Indels:	34
DB:	22
Gaps:	12
US-10-724-972A-6352 (1-335) x US-10-967-189-15 (1-999)	

```

; Publication No. US20030153733A1
; GENERAL INFORMATION:
; APPLICANT: Simpson et al.
; TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB461USD1
; CURRENT APPLICATION NUMBER: US/10/278,946
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: US 09/830,217
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: PCT/US99/06199
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: US 60/078,682
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 60/080,296
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: US 60/084,674
; PRIOR FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 999
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-278-946-15

Alignment Scores:
Pred. No.: 3,28e-25 Length: 999
Score: 405.00 Matches: 113
Percent Similarity: 51.74% Conservative: 65
Best Local Similarity: 32.88% Mismatches: 132
Query Match: 23.96% Indels: 34
DB: 17 Gaps: 12

```

Qy	8	LeuValIleLeuSerValIleGlyLeuLeuPheValLeuIleAlaThrAlaAalaCysGly	27
Db	19	ATTAAATAGCTT--GTTGTACCGTTCCTTCTACTTGTT--TTAGCAGGATGTAGT	72
Qy	28	AsnAenSerSerAsnSerSerysGluSerSerysAspGlyValGluIleIleHis	47
Db	73	GGGAATTCAAATAACAATCATCTGATTAACAAGATAAGSAACAACATCTCAATTAACACAT	132
Qy	48	GluGluGlyThrThrLysValProLysHisProLysArgValValValLeuGluTyrsSer	67
Db	133	GCAATGGGTACAACTGAAATTAAGGGNAACCAAGCGTGTGTACGCATATATCAAGGT	192
Qy	68	PheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsnLys	87
Db	193	GCCACTGACGTCGCTGATCTTAGTGCTTAAACCCTGTAGTGTGTAGTAATCATGACAA	252
Qy	88	LysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyThrSerValGlyThr	107
Db	253	CAAAAACCGAAATTCGAATCATAAAAAATGATTTAAAGAATATAAGATTGTAGGTCAA	312
Qy	108	ArgLysGlnProAenLeuGluGluIleSerLysLeuLysProAspLeuIleAlaAsp	127
Db	313	GAACCTGCACCTAACTTAGAGGAATCTCTAAATTAACAACGGCACTTAATGTGCGCTCA	372
Qy	128	AsnAsnArgHisGlyIleTyLysAspLeuAsnLysIleAlaProThrIleGluLeu	147
Db	373	AAGTTAGAANTGMAAAGTTTACGATCAATTTATCTAAATTCGCACCAACAGTT-----	426
Qy	148	LysSerPheAspGlyAspTyArgnGluAsnIleAspAlaPheLysThrIleSerLysAla	167
Db	427	--TCTACTGATACAGTTTTTCAAAATTCAAA--GATACAACTAAGTTAATGGGGGAAGCT	480
Qy	168	LeuGlyLysGluGluGlyLysLysArgLeuGluHisAspLysLysIleGluGlu	187
Db	481	TTAGGGAACAAAGAAAGCTGAAGATTCTTAAAGTAGCATGATAAAGTAGCTGCA	540
Qy	188	TyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProLaValAlaLys	207



```

355 GAACCTGCACCTAATCTAGAGAAATCTCTAAATATAAACCGGACTTAAATTTGTCGGTCA 414
128 AsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeu 147
415 AAAGTTAGAAATGAAAGTTTACGATCAATATCTAAATCGCACCAACAGTT-----468
148 LysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAla 167
469 ---TCTACTGATACAGTTTTCAAATTCAAA---GATACAACTAAGTTAAATGGGAAAGCT 522
168 LeuGlyLysGluGluGluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGlu 187
523 TTAGGGAAGAAAGAAAGCTGAGATTCTTAAAGATACGATGATATAAGTAGCTGCA 582
188 TyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLys 207
583 TTCAAAAA-----GATGCAAAAGCAAGTATAAGATGATGCCATTGAAA 630
208 SerGlyLeuLeuAlaHisProSerAsn-----SerTyrValGly 220
631 GCTTCAGTTGTTAACTTCCTGCTGATCATACAGAAATTTATGCTGGTGATGCTGGT 690
221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu 240
691 GAAATCTTAATGATTTAGGATTCAAACGT-----AATAAGACTTA 732
241 SerLysTyrLeuLysGly-----ProTyrLeuGlnMetAsnThr---GluThrLeuSer 257
733 CAAAAACAAAGTTGATAATGTTAAGATATTTATCCAACTTACATCTAAGAAACATCCA 792
258 GlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSer 277
793 TTAATGAACCTGATCATATTTTGTAGTAAATCAGATCCAAATCGAAAGATGCTGCA 852
278 Leu-----LysGluLeuGluLysAspProValThrLysLysLysLeuAsnAla 292
853 TTAGTTAAAGAACTGAAAGCGAAGTCACTTCAAGTAAAGAGTGGGAAAAATTTAGACGCA 912
293 ValLysAsnGlnArgVal---AspIleLeuAspArgAspLeuThrAlaArgSerArgGly 311
913 GTTAAACCAACCAAGTATCTGATGATTTAGATGAATCCTTGGAACTTACCTGCGCGA 972
312 LeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysLysSerLys 331
973 TATAATCTTCAATTAAGAACTTTATGACGATTTATATGAA-----AAGTTAAATATTGAA 1026
332 LysAspAsnLys 335
1027 AAACAATCAAAA 1038

```

```

RESULT 24
US-10-329-624-238
; Sequence 238, Application US/10329624
; Publication No. US20040043037A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
;           Gil H. Choi
;           Patrick S. Dillon
;           Craig A. Rosen
;           Steven C. Barash
;           Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

```

```

; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/329,624
; FILING DATE: 27-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/956,171
; FILING DATE: October 20, 1997
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 238:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3775 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 238:
US-10-329-624-238

```

```

Alignment Scores:
Pred. No.: 1-67e-24 Length: 3775
Score: 405.00 Matches: 113
Percent Similarity: 51.74% Conservative: 65
Best Local Similarity: 32.85% Mismatches: 132
Query Match: 23.96% Indels: 34
DB: 19 Gaps: 12

US-10-724-972A-6352 (1-335) x US-10-329-624-238 (1-3775)
QY 8 LeuLysIleLeuSerValIleGlyLeuPheValLeuIleAlaThrAlaLaCysGly 27
Db 61 ATTAATAATGCTT---GTTGTTACGCTTCTTCTACTTGT---TTACGAGATGTAGT 114
QY 28 AsnAsnSerSerSerAsnSerSerLysGluSerLysLysGlyValGluLysHis 47
Db 115 GGAATTTCAATAAACAATCATCTGATAACAAAGATAAGAAACAACATTTCAATTAACAT 174
QY 48 GluGluGlyThrThrLysValProLysHisProLysArgValValValValLeuGluTyrSer 67
Db 175 GCAATGGGTACAACTGAAATTAAGGGGAAACCAAGCGTGTGTTACCGCTATATCAAGGT 234
QY 68 PheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLys 87
Db 235 GCCACTGACGTCGCTGATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGATCATGGACA 294
QY 88 LysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThr 107
Db 295 CAAAACCGAAATTCGAATACATATAAATAATGATTTAAAGATATACTAAGATTTGATGTC 354
QY 108 ArgLysGlnProAsnLeuGluLysSerLysLysLysProAspLeuIleAlaAsp 127
Db 355 GAACCTGCACCTAATCTAGAGAAATCTCTAAATTAACACCGGACTTAAATTTGTCGCTCA 414
QY 128 AsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeu 147
Db 415 AAAGTTAGAAATGAAAGTTTACGATCAATATCTAAATCGCACCAACAGTT-----468
QY 148 LysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAla 167
Db 469 ---TCTACTGATACAGTTTTCAAATTCAAA---GATACAACTAAGTTAAATGGGAAAGCT 522
QY 168 LeuGlyLysGluGluGluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGlu 187

```

523 TTAGGGAAGAAAAAGAGCTGAAGATTACTTTAAAAAGTACGATGATAAAGTAGTGCA 582  
188 TyLysLysGluLeuThrMetAspLysAsnGlnLysValLeuProAlaValAlaLys 207  
583 TTCCAAAA-----GATGCAAAAGCAAGATATAAAGATGATGCCCATTTGAAA 630  
208 SerGlyLeuLeuAlaHisProSerAsn-----SerTyValGly 220  
631 GCTTCAGTTGTTAACTCCGCTGCTGATCATCAAGAATTATGCTGGTGATATGCTGGT 690  
221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu 240  
691 GAAATCTTAAATGATTAGGATTCAAACGT-----AATAAAGACTTA 732  
241 SerLysTyLysLysGly-----ProTyLeuGlnMetAsnThr-----GluThrLeuSer 257  
733 CAAAAACAAGTTAATAATGGTAAAGATATTATCCAACTTACATCTAAAGAAAGCATTTCCA 792  
258 GlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSerAsnGluProSer 277  
793 TTAATGAACGCTGATCATATTTTGTAGTAAATCAGATCCAAATGCGAAAGATGCTGCA 852  
278 Leu-----LysGluLeuGluLysAspProValTrpLysLysLeuAsnAla 292  
853 TTAGTTAAAAAGACTGAAGCGAATGGACTTCAAGTAAAGAGTGGAAAAATTTAGACGCA 912  
293 ValLysAsnGlnArgVal---AspIleLeuAspArgAspLeuTrpAlaArgSerArgGly 311  
913 GTTAAAAACAACCAAGTATCTGATGATTAGATGAAATCATTGGAACTTAGTGGCGGA 972  
312 LeuLeuSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAspSerLys 331  
973 TATAAATCTTCAATTAATAACTTTATGACGATTATATGAA-----AAGTTAAATATTGAA 1026  
332 LysAspAsnLys 335  
1027 AAACAATCAAAA 1038

RESULT 25  
US-10-282-122A-9415  
; Sequence 9415, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Lianguo  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 9415  
; LENGTH: 972  
; TYPE: DNA  
; ORGANISM: Bacillus anthracis  
US-10-282-122A-9415

Alignment Scores:  
Pred. No.: 5,06e-24 Length: 972  
Score: 391.00 Matches: 116  
Percent Similarity: 51.93% Conservative: 59  
Best Local Similarity: 34.42% Mismatches: 122  
Query Match: 23.14% Indels: 40  
DB: 13 Gaps: 13

US-10-724-972A-6352 (1-335) x US-10-282-122A-9415 (1-972)

Qy 9 LysIleLeuSerValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCysGlyAsn 28  
Db 7 AAAATTCCTCAGTATT-----TTCATAGTAGTATTTTCTATTCGCTGTTGGATGC 54  
Qy 29 AsnSerSerSerSerSerSerLysGluSerSerLysAsp-----GlyValGlu 44  
Db 55 GGACAGCAAAAAGAGGAGAGAAAAAGAAACCAAAACCGGACCAATAAATAACAGCTATAACA 114  
Qy 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64  
Db 115 ATTAAACACGCTGAAGGGGAAACGAAGTTAGATAAACCCAGCAAAAGAAAGTAGTTGTTACTT 174  
Qy 65 GluTySerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84  
Db 175 GAATGGGTATATTCAGAAGACTTATTAGCACTTCGTTTCAGCCAGTAGGAGTGGCAGAC 234  
Qy 85 AspAsnLysLysAsnArgIleIle-----LysProLeuArgAspLysIleGly 100  
Db 235 ATTAAGATTATAATAATGGTAAATACAAAAACAAACCCAGTAAAGATGTTGTA--- 291  
Qy 101 LysTyThrSerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLysLys 120  
Db 292 -----GATGTCGGCAGCAGTCAACCAACCACTTAGAAGAAATTTAGCGGTTTAAAAA 342  
Qy 121 ProAspLeuIleAlaAspAsnAsnArgHisLysGlyIleTyLysAspLeuAsnLys 140  
Db 343 CCAGATTTAATTATCACAGCTTCATTCGCTGTTAAAGCAATTTAAAAATGAATTAGAACAA 402  
Qy 141 IleAlaProThrIleGluLeu-----LysSerPheAspGlyAspTyAsnGluAsn 157  
Db 403 ATTGACCAACACAGTTATGTTGATCCATCAACACCAATTAACGATCATTCTGCTGAATG 462  
Qy 158 IleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGluGlyLysArg 177  
Db 463 ACAGAAACATTTTAAACAATTCGAAAGCAGTTGCAAAAGAGAGAGAGAGTAAAAAGTA 522  
Qy 178 LeuGluGluHisAspLysLysIleGluGluTyLysLysGluIle-----Thr 193  
Db 523 TTAGCTGATATGATAAAGCATTCGCTGATGCAAAAGCAAAATTTGAAAAAGCAGACTTA 582  
Qy 194 MetAspLysAsnGlnLysValLeuProAlaValAlaLysSer-----Gly 209  
Db 583 AAAGATAAAAAACATCGCAATGGCAGACAGCATTTACTGCTTAAAAATGTCGCAACATTCG 642  
Qy 210 LeuLeuAlaHisProSerAsnSerTyValGlyGlnPheLeuSerGlnLeuGlyPheLys 229  
Db 643 ATCTTAACT-----GACAAATCTTTAGCTTTTACAAAGTTTACAAAAAATTTAGGT----- 690

QY	230	GluAlaLeuSerAspValThrLysGlyLeuSerLysTyrlLeuLysGlyProTyrLeu	249
Dd	691	:::TTAACAAATCTTTGAAGCAGGAAATCTCGA-----CCTGATGGTTTCAAA	738
QY	250	GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg---MetPheIleMetThr	268
Dd	739	:::     :::     :::     :::     :::	798
QY	269	AsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGluLysAspProValTrpLys	288
Dd	799	:::     :::     :::     :::     :::	849
QY	289	LysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArg	308
Dd	850	:::     :::     :::     :::     :::	909
QY	309	SerArgGglyLeuIleSerSerGluGluMetalLysGluLeuValGluLeu	325
Dd	910	GGTCTTGAGTCGTGCACATCTTTTAGCACACCAAGTAGCAGATGTA	957

[illegible]

•

Publication No. US20050196754A1  
GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc  
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
FILE REFERENCE: 790CIP3/US  
CURRENT APPLICATION NUMBER: US/10/450,763  
CURRENT FILING DATE: 2003-06-11  
PRIOR APPLICATION NUMBER: PCT/US01/08631  
PRIOR FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: 09/540,217  
PRIOR FILING DATE: 2000-03-31  
PRIOR APPLICATION NUMBER: 09/649,167  
PRIOR FILING DATE: 2000-08-23  
NUMBER OF SEQ ID NOS: 60736  
SOFTWARE: Custom  
SEQ ID NO 842  
LENGTH: 4392  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SIMILAR  
LOCATION: (1)-(654)  
OTHER INFORMATION: 98% homologous to Escherichia coli K12 citrate-dependent iron  
OTHER INFORMATION: transport, periplasmic protein, accession number AE000499, Smith-  
OTHER INFORMATION: Waterman Score=1060.  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)-(4392)  
OTHER INFORMATION: n = a,t,c or g  
US-10-450-763-8842  
Alignment Scores:  
Pred. No.: 7.58e-22 Length: 4392  
Score: 375.00 Matches: 96  
Percent Similarity: 50.17% Conservative: 52  
Best Local Similarity: 32.54% Mismatches: 115  
Query Match: 22.19% Indels: 32  
DB: Gaps: 5  
US-10-724-972A-6352 (1-335) x US-10-450-763-8842 (1-4392)  
Qy 45 IlelyshisGluGluGlyThrThrLysValProlyshisProlyshisArgValValValLeu 64  
Db 70 GTTCAGGACGAAACGCGCATCTTAACTCGAAACACGCGCAACGCGATTTGGTGTG 129  
Qy 65 GlutyrSerPheValAspAlaLeuValAlaLeuAspVallyshisProvalGlyIleAlaAsp 84  
Db 130 GAATCTCTGTTCCGCGATGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 189  
Qy 85 AspAenlyshisAsnArgIleIlelyshisProleuArgAspLysIleGlyLysThrSer 104  
Db 190 GATAACGATGCAAAACGCGCATCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 249  
Qy 105 ValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIle 124  
Db 250 GTGCGAACG 309  
Qy 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144  
Db 310 ATTGCGGACGACGATGCGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 369  
Qy 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164  
Db 370 CTGCTGCTTAAGTCTCCCGCAACGAAACCTACGCTGAAATTTTGCAATCTCGCGCTATCATC 429  
Qy 165 SerLysAlaLeuGlyLysGluGluGluGlyLysArgLeuGluGluLysAspLysLys 184  
Db 430 GCGGAAATGGTGGTAAAGGACGAGATGCGAGGACGCGCTCTGGAACAACTTAAGAGAGG 489  
Qy 185 IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204  
Db 490 ATGGCGCATGGGCG 543

Alignment Scores:  
Pred. No.: 3.27e-22 Length: 2209  
Score: 375.00 Matches: 96  
Percent Similarity: 50.17% Conservative: 52  
Best Local Similarity: 32.54% Mismatches: 115  
Query Match: 22.19% Indels: 32  
DB: Gaps: 5  
US-10-724-972A-6352 (1-335) x US-10-450-763-7183 (1-2209)  
Qy 45 IlelyshisGluGluGlyThrThrLysValProlyshisProlyshisArgValValValLeu 64  
Db 2140 GTTCAGGACGAAACGCGCATCTTAACTCGAAACACGCGCAACGCGATTTGGTGTG 2081  
Qy 65 GlutyrSerPheValAspAlaLeuValAlaLeuAspVallyshisProvalGlyIleAlaAsp 84  
Db 2080 GAATCTCTGTTCCGCGATGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2021  
Qy 85 AspAenlyshisAsnArgIleIlelyshisProleuArgAspLysIleGlyLysThrSer 104  
Db 2020 GATAACGATGCAAAACGCGCATCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1961  
Qy 105 ValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIle 124  
Db 1960 GTGCGAACG 1901  
Qy 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144  
Db 1900 ATTGCGGACGACGATGCGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1841  
Qy 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164  
Db 1840 CTGCTGCTTAAGTCTCCCGCAACGCGCATCTGCGCGCGCGCGCGCGCGCGCGCGCG 1781  
Qy 165 SerLysAlaLeuGlyLysGluGluGlyLysArgLeuGluGluLysAspLysLys 184  
Db 1780 GCGGAAATGGTGGTAAAGGACGAGATGCGAGGACGCGCTCTGGAACAACTTAAGAGAGG 1721  
Qy 185 IleGluGluTyrLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204  
Db 1720 ATGGCGCGATGGCG 1667  
Qy 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224  
Db 1666 TCACGCGAACAGCAATTAACCTGCGATCTCAGGAGACCTGGACCGCGCGCGCGCGCG 1607  
Qy 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLysSerLysTyrLeu 244  
Db 1606 TCTCTGGGCGCTG-----AACGTTCCGCGCGCGATG 1577  
Qy 245 LysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMet 264  
Db 1576 GCGGGTGGCTCATGCGCGCTCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1517  
Qy 265 PheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLysAsp 284  
Db 1516 TCTATCCATGCAATTAACGCTTATACATCAACACACCGCTGCTC----- 1475  
Qy 285 ProValTyrLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeu----- 301  
Db 1474 -----TGGACGCGCGGGTCTGTGATCGCGTCCGCTCGCGTTCGCGCTCTGAAATGACC 1421  
Qy 302 -----AspArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGlu 316  
Db 1420 GACGAAGCGGTGAAGATGCGAAACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1361  
Qy 317 -----GluMetAlaLysGluLeuValGluLeuSerLys 327  
Db 1360 TAGCATCACATTAAAGACGTTAACGATCTGCTGGAACCTGCTGAA 1316  
RESULT 28  
US-10-450-763-8842  
; Sequence 8842, Application US/10450763



QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224  
Db 544 TCACGCCAAGCAGCAATTCACCTGATCATCAGGAGACCTGGACCGGACGGTGTGGCC 603  
QY 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu 244  
Db 604 TCTCTGGGGCTG-----AACGTTCCCGCTCCGATG 633  
QY 245 LysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMet 264  
Db 634 GCGGGTGGCTCCATCGCTCTCGCTGGTCCGCTGGCGGCTTCATGTCCACCGGATGT 693  
QY 265 PheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGluLysAsp 284  
Db 694 TCTATCCATGCAATTACGGTTACATCAACACACACCTGTCTC----- 735  
QY 285 ProValTyrLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeu----- 301  
Db 736 -----TGGACGGCGGGTCTGTGATCCGTTCGGTCCGGTTCGGTTCGAAATGACC 789  
QY 302 -----AspArgAspLeuTyrAlaArgSerArgGlyLeuLysSerGlu 316  
Db 790 GACGAAGCGGTGAAGATGCAAACTGGTTGCGGTTCCGCACACGACGCTGACCAAGAA 849  
QY 317 -----GluMetAlaLysGluLeuValGluLeuSerLys 327  
Db 850 TAGGATCATTTAAAGACGTTACGATCTGCCTGAACTGCTGAAA 894

RESULT 29  
US-10-450-763-13139  
; Sequence 13139, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/549,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 13139  
; LENGTH: 2799  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIMILAR  
; LOCATION: (181)..(903)  
; OTHER INFORMATION: 73% homologous to Mus musculus Pro-Pol-dutPase  
; OTHER INFORMATION: polyprotein, accession number Y12713.Smith-Waterman Score=892.  
US-10-450-763-13139

Alignment Scores:  
Pred. No.: 6.14e-20 Length: 2799  
Score: 350.00 Matches: 69  
Percent Similarity: 68.92% Conservative: 33  
Best Local Similarity: 46.62% Mismatches: 46  
Query Match: 20.71% Indels: 0  
DB: 24 Gaps: 0

US-10-724-972A-6352 (1-335) x US-10-450-763-13139 (1-2799)

QY 45 IleLysHisGluGluGlyThrLysValProLysHisProLysArgValValLeu 64  
Db 1672 GTTCAGGACCAACGCGCATGTTTACATCTCAAAAAACGCCACACGGATTGTGTGCTG 1731  
QY 65 GluTyr-SerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84

Db	1732	GAAC	TCTCGTTCGGCCGATGCGCTGGCCGCGCGTGGACGCTCATCCCGATCGGTATTGCGCGAC	1791
QY	85	AspAsn	LysLysAsnAsnTgillellyPrProLeuArgAspLyslleGlyLysTyrThrSer	104
Db	1792	GATA	ACGATGCAAAACGATCTCTGCCGAGTGGCTGGCAGCTTGAACACCGTGGCAGTCC	1851
QY	105	ValGly	ThrArgLysGlnProAsnLeuGluGluileSerLysLeuLysProAspLeuile	124
Db	1852	GTTC	GAACGCGCGCGAGCCGAGCCTTGAAGCCATTCCTCCGCTCTGAAACCCAGACCTGATC	1911
QY	125	IleAla	AspAsnAsnAsnArgHisLysGlylleTyrLysAspLeuAsnLyslleAlaProThr	144
Db	1912	ATTG	CCGACACAGTCGCGCATGCGGGGGTTTACATCGCTTGCAGCAAAATCGCGCGGTA	1971
QY	145	IleGlu	LeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle	164
Db	1972	CTGT	CTGCTTAAGTCCCGACGAACCTACGCTGAAATTTGGAAATCTGGCGGTATCATC	2031
QY	165	SerLys	AlaLeuGlyLysGluGluGluGlyLysArgLeuGluGluHisAspLysLys	184
Db	2032	GCG	AAATATGGTGGTAAAAAACGAGAGATGCAGGCAGCTCTGCAACAACATAAGAGAGG	2091
QY	185	IleGlu	GluTyrLysLysGluile	192
Db	2092	ATGG	CGCAGTGGCCAGCCAGCTT	2115
RESULT 30				
US-08-781-986A-604				
; Sequence 604, Application US/08781986A				
; Publication No US20030054436A1				
; GENERAL INFORMATION:				
; APPLICANT: Charles Kunsch				
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences				
; NUMBER OF SEQUENCES: 5255				
; CORRESPONDENCE ADDRESS:				
; ADDRESSEE: Human Genome Sciences, Inc.				
; STREET: 9410 Key West Avenue				
; CITY: Rockville				
; STATE: Maryland				
; COUNTRY: USA				
; ZIP: 20850				
; COMPUTER READABLE FORM:				
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage				
; COMPUTER: HP Vectra 486/33				
; OPERATING SYSTEM: MSDOS version 6.2				
; SOFTWARE: ASCII Text				
; CURRENT APPLICATION DATA:				
; APPLICATION NUMBER: US/08/781,986A				
; FILING DATE:				
; CLASSIFICATION: 435				
; PRIOR APPLICATION DATA:				
; APPLICATION NUMBER:				
; FILING DATE:				
; ATTORNEY/AGENT INFORMATION:				
; NAME: Benson, Bob				
; REGISTRATION NUMBER: 30,446				
; REFERENCE/DOCKET NUMBER: PB248PP				
; TELECOMMUNICATION INFORMATION:				
; TELEPHONE: (301) 309-8504				
; TELEFAX: (301) 309-8512				
; INFORMATION FOR SEQ ID NO: 604:				
; SEQUENCE CHARACTERISTICS:				
; LENGTH: 2115 base pairs				
; TYPE: nucleic acid				
; STRANDEDNESS: double				
; TOPOLOGY: linear				
US-08-781-986A-604				
Alignment Scores:				
Pred. No.: 1,03e-18 Length: 2115				
Score: 334.00 Matches: 67				
Percent Similarity: 63.84% Conservative: 46				
Best Local Similarity: 37.85% Mismatches: 58				

Query Match: 19.76% Indels: 6  
DB: 8 Gaps: 3

US-10-724-972A-6352 (1-335) x US-08-781-986A-604 (1-2115)

QY 154 TyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 173  
DB 3 TATAATGCAATATTGAAGCATTTAAACACAGTCGCTAAAGCAGTAGGCAAGAGAA 62

QY 174 GlyLysArgLeuGluGluHisAspLysLysLysLysLysLysLysLysLysLys 193  
DB 63 GGCAGAGAGCGCTCGAAAGCATGATAAATATTAGCGGAGATTAGAAAGAAATTTGAA 122

QY 194 MetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLysLysLys 213  
DB 123 CAGAGTAGCTTAAATCTGCATTTGCATTTGCATTTGCATTTGCATTTGCATTT 182

QY 214 ProSerAsnSerTyrValGlyGlnPheLysGlnLeuGlyPheLysGluAlaLeuSer 233  
DB 183 AATGAAGATACATTTATGGGACAAATCTTAATTAATGAGTATTTCAACCTGAGTCAMA 242

QY 234 AspAspValThrLysGlyLysLysLysLysLysLysLysLysLysLysLysLys 253  
DB 243 AAARAMAAACTACGCAATGTTGGTGAACGCAAGCGTGTCTTATATATATATTTAAATAAT 302

QY 254 GluThrLeuSerGlnValAsnProGluArgMetPheLysMetThrAsn---LysAlaSer 272  
DB 303 GAAGCACTTGCCAAATCAATCCAAAGTATGATTTTGCACCTGACGGAACAAACGGAC 362

QY 273 SerAsnGluProSerLysGlyLysLysLysLysLysLysLysLysLysLysLys 291  
DB 363 AAAAATAGAACGAAATTC-----ATTGATCTCGAGTTTGGAAATCATTTAAA 410

QY 292 AlaValLysAsnGlnArgValAspLysLysLysLysLysLysLysLysLysLys 311  
DB 411 GCTGTGAAGATAACAAAGTTTATGACGTTGACCGAATAAGTGTTCGAAATCAAGGGG 470

QY 312 LeuLysSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328  
DB 471 ATTATCGCAAGTGAAGTATGCGAAGATTTAGAAAAAATTTGCAGAAAAA 521

RESULT 31

US-10-329-624-604  
; Sequence 604, Application US/10329624  
; Publication No. US20040043037A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; Gil H. Choi  
; Patrick S. Dillon  
; Craig A. Rosen  
; Steven C. Barash  
; Michael R. Rannon  
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5256  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/329,624  
; FILING DATE: 27-Dec-2002  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/956,171  
; FILING DATE: October 20, 1997  
; APPLICATION NUMBER: 60/009,861

FILING DATE: January 5, 1996  
APPLICATION NUMBER: 08/781,986  
FILING DATE: January 3, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mark J. Hyman  
REGISTRATION NUMBER: 46,789  
REFERENCE/DOCKET NUMBER: PB248P1D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (240) 314-1224  
TELEFAX: (301) 309-8439  
INFORMATION FOR SEQ ID NO: 604:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2115 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 604:  
US-10-329-624-604

Alignment Scores:  
Pred. No.: 1,03e-18 Length: 2115  
Score: 334.00 Matches: 67  
Percent Similarity: 63.84% Conservative: 46  
Best Local Similarity: 37.85% Mismatches: 58  
Query Match: 19.76% Indels: 6  
DB: 19 Gaps: 3

US-10-724-972A-6352 (1-335) x US-10-329-624-604 (1-2115)

QY 154 TyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 173  
DB 3 TATAATGCAATATTGAAGCATTTAAACACAGTCGCTAAAGCAGTAGGCAAGAGAA 62

QY 174 GlyLysArgLeuGluGluHisAspLysLysLysLysLysLysLysLysLysLys 193  
DB 63 GGCAGAGAGCGCTCGAAAGCATGATAAATATTAGCGGAGATTAGAAAGAAATTTGAA 122

QY 194 MetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLysLysLys 213  
DB 123 CAGAGTAGCTTAAATCTGCATTTGCATTTGCATTTGCATTTGCATTTGCATTT 182

QY 214 ProSerAsnSerTyrValGlyGlnPheLysGlnLeuGlyPheLysGluAlaLeuSer 233  
DB 183 AATGAAGATACATTTATGGGACAAATCTTAATTAATGAGTATTTCAACCTGAGTCAMA 242

QY 234 AspAspValThrLysGlyLysLysLysLysLysLysLysLysLysLysLysLys 253  
DB 243 AAARAMAAACTACGCAATGTTGGTGAACGCAAGCGTGTCTTATATATATATTTAAATAAT 302

QY 254 GluThrLeuSerGlnValAsnProGluArgMetPheLysMetThrAsn---LysAlaSer 272  
DB 303 GAAGCACTTGCCAAATCAATCCAAAGTATGATTTTGCACCTGACGGAACAAACGGAC 362

QY 273 SerAsnGluProSerLysGlyLysLysLysLysLysLysLysLysLysLysLys 291  
DB 363 AAAAATAGAACGAAATTC-----ATTGATCTCGAGTTTGGAAATCATTTAAA 410

QY 292 AlaValLysAsnGlnArgValAspLysLysLysLysLysLysLysLysLysLys 311  
DB 411 GCTGTGAAGATAACAAAGTTTATGACGTTGACCGAATAAGTGTTCGAAATCAAGGGG 470

QY 312 LeuLysSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328  
DB 471 ATTATCGCAAGTGAAGTATGCGAAGATTTAGAAAAAATTTGCAGAAAAA 521

RESULT 32

US-08-781-986A-2556  
; Sequence 2556, Application US/08781986A  
; Publication No. US20030054436A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5255

;;  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Human Genome Sciences, Inc.  
;; STREET: 9410 Key West Avenue  
;; CITY: Rockville  
;; STATE: Maryland  
;; COUNTRY: USA  
;; ZIP: 20850  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
;; COMPUTER: HP Vectra 486/33  
;; OPERATING SYSTEM: MSDOS version 6.2  
;; SOFTWARE: ASCII Text  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/781,986A  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Benson, Bob  
;; REGISTRATION NUMBER: 30,446  
;; REFERENCE/DOCKET NUMBER: PB248PP  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (301) 309-8504  
;; TELEFAX: (301) 309-8512  
;; INFORMATION FOR SEQ ID NO: 2556:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 242 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear  
;;  
US-08-781-986A-2556  
Alignment Scores:  
Pred. No.: 8.86e-20 Length: 242  
Score: 333.00 Matches: 66  
Percent Similarity: 90.00% Conservative: 6  
Best Local Similarity: 82.50% Mismatches: 8  
Query Match: 19.70% Indels: 0  
DB: Gaps: 0  
US-10-724-972A-6352 (1-335) x US-08-781-986A-2556 (1-242)  
QY 70 AspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsnLysLysAsn 89  
Db 2 GATGCATTAGCAGCATTAGACGTTAAACCCAGTTGGTATTGCTGATGATGTAAGAAAAA 61  
QY 90 ArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLys 109  
Db 62 CGTATCATTAACCCAGTTAGAGAAAAATTGGGATTATTAATTCCTGATGATGTAAGAAAAA 121  
QY 110 GlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsnAsn 129  
Db 122 CAGCAAACTTAGAAGAAATTAGTAAATTAACCCGATTATTAATTCCTGATGATGTAAGAAAAA 121  
QY 130 ArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSer 149  
Db 182 AGACATANAGGTATTAAATAAGAAATTAAACAAATTAACCAAAATTCACCAACATTATCATTAAGAGT 241  
RESULT 33  
US-10-329-624-2556  
; Sequence 2556, Application US/10329624  
; Publication No. US2004004037A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; Gil H. Choi  
; Patrick S. Dillon  
; Craig A. Rosen  
; Steven C. Barash  
; Michael R. Fannon  
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5256

;;  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Human Genome Sciences, Inc.  
;; STREET: 9410 Key West Avenue  
;; CITY: Rockville  
;; STATE: Maryland  
;; COUNTRY: USA  
;; ZIP: 20850  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
;; COMPUTER: HP Vectra 486/33  
;; OPERATING SYSTEM: MSDOS version 6.2  
;; SOFTWARE: ASCII Text  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/10/329,624  
;; FILING DATE: 27-Dec-2002  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/956,171  
;; FILING DATE: October 20, 1997  
;; APPLICATION NUMBER: 60/009,861  
;; FILING DATE: January 5, 1996  
;; APPLICATION NUMBER: 08/781,986  
;; FILING DATE: January 3, 1997  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Mark J. Hyman  
;; REGISTRATION NUMBER: 46,789  
;; REFERENCE/DOCKET NUMBER: PB248PID1  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (240) 314-1224  
;; TELEFAX: (301) 309-8439  
;; INFORMATION FOR SEQ ID NO: 2556:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 242 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear  
;;  
US-10-329-624-2556  
Alignment Scores:  
Pred. No.: 8.86e-20 Length: 242  
Score: 333.00 Matches: 66  
Percent Similarity: 90.00% Conservative: 6  
Best Local Similarity: 82.50% Mismatches: 8  
Query Match: 19.70% Indels: 0  
DB: Gaps: 0  
US-10-724-972A-6352 (1-335) x US-10-329-624-2556 (1-242)  
QY 70 AspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsnLysLysAsn 89  
Db 2 GATGCATTAGCAGCATTAGACGTTAAACCCAGTTGGTATTGCTGATGATGTAAGAAAAA 61  
QY 90 ArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLys 109  
Db 62 CGTATCATTAACCCAGTTAGAGAAAAATTGGGATTATTAATTCCTGATGATGTAAGAAAAA 121  
QY 110 GlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsnAsn 129  
Db 122 CAGCAAACTTAGAAGAAATTAGTAAATTAACCCGATTATTAATTCCTGATGATGTAAGAAAAA 121  
QY 130 ArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSer 149  
Db 182 AGACATANAGGTATTAAATAAGAAATTAAACAAATTAACCAAAATTCACCAACATTATCATTAAGAGT 241  
RESULT 34  
US-09-738-626-375  
; Sequence 375, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO

```

/ APPLICANT: OCHIAI, KEIKO
/ APPLICANT: YOKOI, HARUHIKO
/ APPLICANT: TATEISHI, NAKKO
/ APPLICANT: SENOH, AKIHIRO
/ APPLICANT: IKEDA, MASATO
/ APPLICANT: OZAKI, AKIO
/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
/ FILE REFERENCE: 249-125
/ CURRENT APPLICATION NUMBER: US/09/730,626
/ CURRENT FILING DATE: 2000-12-18
/ PRIOR APPLICATION NUMBER: JP 99/377484
/ PRIOR FILING DATE: 1999-12-16
/ PRIOR APPLICATION NUMBER: JP 00/159162
/ PRIOR FILING DATE: 2000-04-07
/ PRIOR APPLICATION NUMBER: JP 00/280988
/ PRIOR FILING DATE: 2000-08-03
/ NUMBER OF SEQ ID NOS: 7059
/ SOFTWARE: PatentIn ver. 3.0
/ SEQ ID NO 375
/ LENGTH: 945
/ TYPE: DNA
/ ORGANISM: Corynebacterium glutamicum
US-09-738-626-375

Alignment Scores:
Pred. No.: 6,97e-19 Length: 945
Score: 331.00 Matches: 90
Percent Similarity: 49.21% Conservative: 65
Best Local Similarity: 28.57% Mismatches: 140
Query Match: 19.59% Indels: 20
DB: 9 Gaps: 7

US-10-724-972A-6352 (1-335) x US-09-738-626-375 (1-945)
Qy 13 ValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCysGlyAsnAsnSerSerSer 32
Db 43 GTGGCCCGGCTCTCGCGGTGGCGACTCTCTTAAGCTCTGTTCTTCAACATCTCCGAT 102
Qy 33 AsnSer---SerLysGluSerSerLysAspGlyValGluIleLysHisGluGluGlyThr 51
Db 103 GAATCAATCCAACTGAAGTTCGCCAGCACTGGATATTCAAGTGGAGCACGCAATGGCCACC 162
Qy 52 ThrLysValProLysHisProLysArgValValValValLeuGluTyrSerPheValAspAla 71
Db 163 ACCGAAATCCCTGAAACCCCAACCGCTGTGTGCTGTCATTGATCCCCACACCTCGACGCA 222
Qy 72 LeuValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLysLysAsnArgIle 91
Db 223 CTTTGGGTTTTGGGAATTAATCTCCAGTCCGAGACTCGGAATCTGGATCCGAAAATGGTTTC 282
Qy 92 IleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnPro 111
Db 283 CCCGCTACTTGGCTGACGAGCTAAAGACACCGAATCTGTTGGGCTGACATCTCGAGCA 342
Qy 112 AsnLeuGluGluIleSerLysLeuLysProAspLeuIleIleAlaAspAsnAsnArgHis 131
Db 343 AATTTGGAAAAGATCGCCGCACTGGATCCGGAATTTGATCATTTGGCGCAAGGTCGCCGAC 402
Qy 132 LysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAsp 151
Db 403 GAGGCTATTATGATCATGCTTTTCAGACATCGGCAACACCGTGATGTCGAAAGTTCCGGC 462
Qy 152 GlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGlu 171
Db 463 ACAAACTGGANTGAACAGCGAGAA-----ATCACTGGCGCAGCAGTAAACAAGTCT 513
Qy 172 GluGluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGluTyrLysLysGlu 191
Db 514 GATGAGATGGACAAACTGATCTCAGACTTTGGACACCCGTCGCCACAGAGCTTGGTGAAGAG 573
Qy 192 IleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeu 211
Db 574 ATCGGTGTCGTCCGCGACAAACCCGCTTCAATGTTGTCGATTCGGCACGACCACTTCAGGCTC 633

```

QY	212	AlaHisProSerAanSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAla	233
Db	634	TATGGTCCC---GAGACCTTCCTCGTTCAATTCTTCGAAACAAGTCGGATTT-----	681
QY	232	LeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMet	251
Db	682	-----GACCCTGGGGGAACGGTGATGGAAATGAGTAC-----TCCATGATGGAGCTA	726
QY	252	AenThrGluThrLeuSerGlnValAasnProGluArgMetPhe-----IleMetThrAan	269
Db	727	TCCTCAGAAAACCTTTGGGCAGATCGATGGAGACTTATTTTCTACACCATCCCGAGGATCC	786
QY	270	LysAlaSerSerAasnGluProSerLeuLysGluLeuLysAspProValTrpLysLys	289
Db	787	CCTGAAGCAACCACTTATCCAAAGATTCCGAACTG-----TGGGTGTGAT	831
QY	290	AasnAlaValLysAasnGlnArgValAaspIleLeuAspArgAepLeuTrpAlaLaArgSer	309
Db	832	TCACCAGCAGTTCGGCAAGGTAGAACCTTACGAGTTTGAAGACGAAACCTCGATGGTCCGC	891
QY	310	ArgGlyLeuIleSerSerGluGluMetAlalysGluLeuValGlu	324
Db	892	ATCGGTGATTAGGTGCCAATGAAATCTTGGATGACCTGGGAAGAA	936
 RESULT 35 US-09-738-626-1 ; Sequence 1, Application US/09738626 ; Publication No. US20020197605A1 ; GENERAL INFORMATION: ; APPLICANT: NAKAGAWA, SATOSHI ; APPLICANT: MIZOGUCHI, HIROSHI ; APPLICANT: ANDO, SEIKO ; APPLICANT: HAYASHI, MIKIRO ; APPLICANT: OCHIAI, KEIKO ; APPLICANT: YOKOI, HARUHIKO ; APPLICANT: TATEISHI, NAOKO ; APPLICANT: SENO, AKIHIO ; APPLICANT: IKEDA, MASATO ; APPLICANT: OKAZKI, AKIO ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES ; FILE REFERENCE: 249-125 ; CURRENT APPLICATION NUMBER: US/09/738,626 ; CURRENT FILING DATE: 2000-12-18 ; PRIOR APPLICATION NUMBER: JP 99/377484 ; PRIOR FILING DATE: 1999-12-16 ; PRIOR APPLICATION NUMBER: JP 00/159162 ; PRIOR FILING DATE: 2000-04-07 ; PRIOR APPLICATION NUMBER: JP 00/280988 ; PRIOR FILING DATE: 2000-08-03 ; NUMBER OF SEQ ID NOS: 7059 ; SOFTWARE: PatentIn ver. 3.0 ; SEQ ID NO 1 ; LENGTH: 3309400 ; TYPE: DNA ; ORGANISM: Corynebacterium glutamicum US-09-738-626-1			
 Alignment Scores:			
Pred. No.:	1.51e-14	Length:	3309400
Score:	331.00	Matches:	90
Percent Similarity:	49.21%	Conservative:	65
Best Local Similarity:	28.57%	Mismatches:	140
Query Match:	19.59%	Indels:	20
DB:	9	Gaps:	7
 US-10-724-972A-6352 (1-335) x US-09-738-626-1 (1-3309400)			
QY	13	VallileGlyLeuLeuPheValLeulleAlaThrAlaAlaCysGlyAasnAanSerSer	324
Db	352735	GTGGCCCCGTCTTCGGCGTCGCACCTCTTAAGACTCTCTTCTTCAACATCTTCCGAT	355
QY	33	AanSer----SerLysGluSerSerLysAaspGlyValGluIleLysHisGluGluGlyThr	51

... 0 504 0503 0250 (1 23E) x 118-09-238-626-1 (1-3309400)

Qy 13 valleclyleuLeuPheValleuLeuAlaAlaCysGlyAsnAsnSerSerSer 32  
||| ||||| ||| :: ::::: || :: |||||  
Db 352735 GTGGCCCGCTCTCTGGCGTGGCACTCTTCTTAAGCTCTCTTCAACACTCTTCGGAT 352794  
||| ||||| ||| :: ::::: || :: |||||



APPLICANT: Ohlsen, Kari  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamanoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282.122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 17083  
LENGTH: 963  
TYPE: DNA  
ORGANISM: Clostridium difficile  
US-10-724-972a-6352 (1-335) x US-10-282-122a-17083 (1-963)

Alignment Scores:  
Pred. No.: 1,578-18 Length: 963  
Score: 327.00 Matches: 99  
Percent Similarity: 48.12% Conservative: 55  
Best Local Similarity: 30.94% Mismatches: 136  
Query Match: 19.35% Indels: 30  
DB: 18 Gaps: 12

US-10-724-972a-6352 (1-335) x US-10-282-122a-17083 (1-963)

QY 11 LeuSerValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCysGlyAsnAsnSer 30  
DB 31 ATAGCTATTATAGGATTAGTACAGTGTGTTT-----GCATTAGGTGGAAGTAAAAAG 81  
QY 31 SerSerAsnSerSerLysGluSerSerLysAspGlyValGluIleLysHisGluGluGly 50  
DB 82 AATGAATCAAAAACATCAGAAAAATTCATAATAATACCAATAAGATAACTATAACTTTGGGA 141  
QY 51 ThrThrLysValProLysHisProLysArgValValValValLeuGluTyrSerPheValAsp 70  
DB 142 GAAACAGATGTAAAAATTAATCTCTAAAGAAAGTAGTAGTATTGATTATTCGGCTTTAGAT 201  
QY 71 AlaLeuValAlaLeuAspValLysPro-----ValGlyIleAlaLeuAspAspAsnLysLys 86  
DB 202 ACATGGATGCATTAGGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 261  
QY 89 AsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArg 108  
DB 262 GCATCTTTAGAAAAA---TATAAGGATAAG-----AAGTATACAGATTTAGGAGGCTTA 312  
QY 109 LysGluProAsnLeuGluGluIleSerLysLeuLysProAspLeuIleIleAlaAspAsn 128

Db 313 AAAGAACCAAGATTAGAGGTATTAAGTCGGCAATCCAGATTAAATTATAATA---AAT 369  
QY 129 AsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLys 148  
Db 370 GGNAGACAAGAAGATTATTTATGAGCAATTTGTCAAAAATAGCACCAACAATAGTACAAGT 429  
QY 149 SerPheAspGlyAspTyrAsnGlu-----AsnIleAspAlaPheLysThrIle 164  
Db 430 AAGATGATGAAGAAGTACTTAGAATCTGTAAAAAATAATATAGAC-----AAAAATA 480  
QY 165 SerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAspLysLys 184  
Db 481 GCAAAAATATTTGGAGTAGAAGAAAAAGCAAAATCAAGAATTTAGTAAATTTGAAAGAAA 540  
QY 185 IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204  
Db 541 ATGAAACATTTGAAATAAAAAAGTAACA---GATAAAAAATTTAAATGCTTTAACTATAATG 597  
QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224  
Db 598 GTAAATGAAGTAAATTTAAGTGTATTTGGTGAAGAATCAAGATTTTAGTATATATATACAAT 657  
QY 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu 244  
Db 658 AGTTTTCGATTT-----GAAATAAAGATATAAATAATAAAGATCAAGT 702  
QY 245 LysGlyProTyrLeuGlnMetAsnThrThrLysSerGlnValAsnProGluArgMet 264  
Db 703 CATGGACAA-----AATATTACTTTTGAATATATATAGCTAAACAAAATCCAGAAAGTTATG 756  
QY 265 PheIleMet-----ThrAsnLysAlaSerSerAsnGluProSerLeuLys 279  
Db 757 TTCGTTATAGATAGAGGAATAGCTACTGGCAGTGATGTAAGAAAGAGCTCAACTGCAAAA 816  
QY 280 GluLeuGluLysAspProValTyrLysLysLeuAsnAlaValLysAsnGlnArgValAsp 299  
Db 817 TCTGTTTGAACAATGATATAATAAATCTATGATGATATAGATATAGATATATAATAATA 876  
QY 300 IleLeuAspArgAspLeuTyr---AlaArgSerArgGlyLeuIleSerSerGluGluMet 318  
Db 877 TATTAGATTCCCAACACATGCTATGTAATGATGCTGTCTTAACATCTTTAAATAAAAATG 936

RESULT 38  
US-10-721-922A-351  
Sequence 351, Application US/10721922A  
Publication No. US20050191732A1  
GENERAL INFORMATION:  
APPLICANT: Pompejus, Markus  
APPLICANT: Kroger, Burkhard  
APPLICANT: Schroder, Hartwig  
APPLICANT: Zelder, Oskar  
APPLICANT: Haberhauer, Gregor  
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS  
FILE REFERENCE: BG1-132CPN  
CURRENT APPLICATION NUMBER: US/10/721,922A  
CURRENT FILING DATE: 2003-11-24  
PRIOR APPLICATION NUMBER: 09/6027777  
PRIOR FILING DATE: 2000-02-23  
PRIOR APPLICATION NUMBER: 09/603124  
PRIOR FILING DATE: 2000-06-23  
PRIOR APPLICATION NUMBER: US 60/141031  
PRIOR FILING DATE: 1999-06-25  
PRIOR APPLICATION NUMBER: US 60/143694  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: US 60/151778  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19931418.7  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19932124.8  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932126.4  
PRIOR FILING DATE: 1999-07-09

Qy 256 LeuSerGlnValAsnProGluArgMetPhe-----IleMetThrAsnLysAlaSerSer 273  
 Db 686 TTGGGCAGATCGATGGAGACCTTATTTCTACACCATCCCAAGGATCCCTCGAAGCAACC 745

	.....   :::	796
D <sub>b</sub>	ACTTATCCAAAGATTTCGCACTG-----TGCGTTGTATCCACGACGGTT	790
Q <sub>y</sub>	LysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArgGlyLeuLle	313
	:::	
D <sub>b</sub>	CGCAAGGTAGNACTTACGAGTTTGGAACGACGAAACCTCGGTGGTGGCATCGGTGTATTA	850
Q <sub>y</sub>	SerSerGluGluMetAlaLysGluLeuValGlu	324

DB 851 GGTCCTGAATCTTGGATGACCTGGAAGAA 883

RESULT 39

US-09-974-300-2044

; Sequence 2044, Application US/09974300

; Patent No. US20020146721A1

; GENERAL INFORMATION:

INVENTOR:

APPLICANT: Clausen, Ib Groth

```

; TITLE OF INVENTION: Expression
;
; FILE REFERENCE: 10085.500-US
;
; CURRENT APPLICATION NUMBER: US/09/974,300
;
; CURRENT FILING DATE: 2001-10-05
;
; PRIOR APPLICATION NUMBER: 03/690,598
;
; PRIOR FILING DATE: 2000-10-06

```

```

: PRIOR FILING DATE: 2001-03-27
:
: NUMBER OF SEQ ID NOS: 8481
:
: SOFTWARE: FastSEQ for Windows Version 4.0
: SEQ ID NO 2044
: LENGTH: 957
:

```

US-09-974-300-2044

Alignment Scores:	5.12e-18	Length:	957
Pred. No.:		Score:	321.00
Percent Similarity:		Matches:	98
			56.64

Best Local Similarity:	29.79%	Mismatches:	131
Query Match:	18.99%	Indels:	32
DB:	9	Gaps:	13

US-10-724-972A-6352 (1-335) x US-09-974-300-2044 (1-957)  
Qy. 9 LysIleLeuSerValIleGlyLeuLeuPheValLeuIleAlaThrAlaCysGlyAsn 28

Db	7	AAATATCAATATTTCTATTTATTTCTCTCTCGCGCTTGGCGACGGGGATGCGGGAAC	66
Qy	29	Asn-----SerSerSerAspSerSerLvsGluSerSerLvsAsnGluValGluTleGlu	45

[illegible]

Db  
127 TACCTCGATCATACATAAGGTG--AAGATCCGGCCGAACAATCGTATTACGGGA 183

Ov  
66 -----TTCGCPHATVATL-----

Db 184 AGCGTGAATCGATGAAGATGCCAAGCTT---CTGGATGTTTCACCAGCGGTGCCATT 240

Db	241	TCCTCTCCGGAAGTTC	CGGATCTGTTA	GAGCA	TACTGACA	AAGCA	-----	291
		:::	:::	:::	:::	:::		
		...	...	...	...	...	...	...

[illegible]

Search completed: November 10, 2005, 00:41:18  
Job time : 1811 secs

**RESULT 40**

; Patent No. US20020061569A1

APPLICANT: Havelbeck

APPLICANT: OHLBELL, KATH E.  
APPLICANT: ZUCKERMAN, JUDITH W.

; APPLICANT: Wall, Daniel

APPLICANT: Carr, Grant J.

APPLICANT: JAMES H. HOWARD  
APPLICANT: JAMES H. HOWARD

**INVENTION OF LITTLE**

FILE REFERENCE: ELITRA.011A

**CURRENT FILING DATE:** 2001-03-

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

;; PRIOR APPLICATION NUMBER: 8072  
;; FILING DATE: 2000-05-22

; PRIOR APPLICATION NUMBER: 60/2

PRIOR APPLICATION NUMBER: 60/2

; PRIOR FILING DATE: 2000  
; PRIOR APPLICATION NUMBER: 60/2

; Patent No. US20020061569A1

APPLICANT: **Hasselbeck, Robert**

APPLICANT: OHLBEN, KATH E.  
 ADDRESS: 7241 13th W  
 CITY: DENVER, CO 80202

; APPLICANT: wall, Daniel

APPLICANT: Carr, Grant J.

APPLICANT: TAMMANGO, ROBERT E.  
APPLICANT: XII. H. Howard

1. TITLE OF INVENTION: IDENTIFICATION OF RESEARCH CONCEPTS

FILE REFERENCE: ELITRA.011A

CURRENT FILING DATE: 2001-03-21

; PRIOR AFFILIATION NUMBER: 007-1527-0-1  
 ; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848  
 ; FILING DATE: 2000 05-22

; PRIOR APPLICATION NUMBER: 60/207,727

PRIOR APPLICATION NUMBER: 60/242,578

1. PRIOR FILING DATE: 2000-01-01  
2. PRIOR APPLICATION NUMBER: 60/253,625



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☐ BLACK BORDERS

☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES

☐ FADED TEXT OR DRAWING

☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING

☐ SKEWED/SLANTED IMAGES

☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS

☐ GRAY SCALE DOCUMENTS

☐ LINES OR MARKS ON ORIGINAL DOCUMENT

☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

**This Page Blank (uspto)**